JPRS L/10076

27 October 1981

Translation

ECONOMIC FORECASTING FOR THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

By

S.A. Sarkisyan, et al.

FBIS FOREIGN BROADCAST INFORMATION SERVICE

NOTE

JPRS publications contain information primarily from foreign newspapers, periodicals and books, but also from news agency transmissions and broadcasts. Materials from foreign-language sources are translated; those from English-language sources are transcribed or reprinted, with the original phrasing and other characteristics retained.

Headlines, editorial reports, and material enclosed in brackets [] are supplied by JPRS. Processing indicators such as [Text] or [Excerpt] in the first line of each item, or following the last line of a brief, indicate how the original information was processed. Where no processing indicator is given, the information was summarized or extracted.

Unfamiliar names rendered phonetically or transliterated are enclosed in parentheses. Words or names preceded by a question mark and enclosed in parentheses were not clear in the original but have been supplied as appropriate in context. Other unattributed parenthetical notes within the body of an item originate with the source. Times within items are as given by source.

The contents of this publication in no way represent the policies, views or attitudes of the U.S. Government.

JPRS L/10076

27 October 1981

ECONOMIC FORECASTING FOR THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

Moscow EKONOMICHESKOYE PROGNOZIROVANIYE RAZVITIYA BOL'SHIKH TEKHNICHESKIKH SYSTEM in Russian 1977 (signed to press 10 Jun 77) pp 1-318

[Book by S.A. Sarkisyan, D.E. Starik, P.L. Akopov, E.S. Minayev and V.I. Kaspin, written under the editorial review of Doctor of Economic Sciences V.A. Lisichkin, Izdatel'stvo Mashinostroyeniye, 3,800 copies, 318 pages]

CONTENTS

CHAPTER 1: SCIENTIFIC AND TECHNICAL PROGRESS AND THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

CHAPTER 2: FORECASTING THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

CHAPTER 3: CRITERIA FOR ESTIMATING EFFECTIVENESS OF LARGE TECHNICAL SYSTEMS

[Annotation] The book examines the basic principles and methods of economic forecasting, as well as the specific features of their application to the class of large technical systems (BTS). The development patterns of technology are shown under the conditions of the present-day scientific and technical revolution, the reasons for the occurrence of large technical systems and their distinguishing features and classification. Also taken up are the basic concepts dealing with systems, their life cycle and the principles of technical and economic analysis. The methods are given for forecasting the development of the BTS, the criteria and methods for assessing the economic effectiveness of the systems, the models and methods of forecasting their values.

The book is designed for scientific workers and engineers whose sphere of professional interests includes the questions of forecasting and an economic assessment of scientific and technical progress. It can also be useful for instructors and students in machine building VUZes.

17 tables, 50 illustrations, and a bibliography of 50 titles.

LIST OF STANDARD TRANSLATIONS

1. taktiko-tekhnicheskiye trebovaniya--tactical-technical requirements
2. opytno-konstruktorskaya razrabotka--prototype design work
3. opytnaya sistema--prototype system
4. ekspluatatsiya--operation
5. seriynoye proizvodstvo--serial production
6. sebestoimost'--production cost
7. tekushchiye zatraty--current expenses
8. kapital'nyye vlozheniya--capital investments
9. stroitel'no-montazhnyye raboty--construction-installation work
10. nauchno-issledovatel'skiye raboty--scientific research
11. avanproyekt--preliminary project, design
12. tekhnicheskoye zadaniye--technical requirement, specification
13. tekhnicheskiye predlozheniya--technical proposals
14. eskiznyy proyekt--draft design
15. prorabotka--study
16. maketirovaniye--mock-up construction
17. tekhnologiya--technology, production method
18. opytnyy obrazets--prototype
19. rabochiye chertezhi--working drawings

FOREWORD

A most important condition for increasing the efficiency of social production and improving product quality, as was pointed out by the 25th CPSU Congress, is an accelerated pace of scientific and technical progress based upon comprehensive programs [1]. The comprehensive programs and the long-range development plans compiled on their basis for the interrelated national economic sectors linked together by the common end result of research, design and production activities, are an effective means for ensuring the planned and time-coordinated development of science, technology and production. The developing of the comprehensive scientific and technical development programs and long-term plans involves the necessity of surmounting a whole series of ambiguities.

Among them one would mention the ambiguities relative to: the development goals; the means and methods of achieving the goals; the resources ensuring development; the total development effectiveness; the comparative effectiveness of possible development areas under the conditions of future resource constraints. The first two types of ambituities can be overcome by special methods based, as a rule, on nonformal (heuristic) and formal (extrapolation) forecast assessments.

The theoretical and practical aspects of forecasting can be found in a large number of articles and monographs published in the USSR and abroad. These studies take up a large range of questions related to the gnoseology and methodology of forecasting, the results of the practical implementation of individual methods are given, and the questions of organizing forecast activities are examined. However, it must be pointed out that many questions in the theory and practice of forecasting still remain debated.

At the same time the existing publications virtually do not deal with the methodological aspects of forecasting the resources which ensure scientific and technical development as well as the questions related to assessing the economic consequences of scientific and technical progress. Active control of scientific and technical progress becomes effective only under conditions where, along with assessing its results, consideration is given to the entire spectrum of resources essential for carrying out one or another direction in scientific and technical development.

The problems of forecasting and assessing economic effectiveness from scientific and technical progress assume particular acuteness in line with the necessity of managing the development processes of large technical systems (BTS).

The BTS are a direct consequence of the present-day scientific and technical revolution which made a start to the age of the conquering of space, the use of atomic and nuclear energy, computers, the automation of production processes and so forth. In being marked by great complexity, the BTS require significant resource outlays on their creation and series production. At the same time the use of the BTS for their specific purpose creates an opportunity of obtaining an economic effect in various spheres of human activity. From this arises the problem of correlating the resources consumed in the various stages of the life cycle of the systems with the effect obtained during the period of their operation. The solving of this problem should correlate two aspects of scientific and technical development for the BTS, namely the technical and economic, and this is essential for drawing up the long-range plans.

The book presented for the reader's consideration attempts to systematize an examination of the methods of choosing alternatives for the development of the BTS from the viewpoint of their integrated technical and economic evaluation. In considering that the BTS properties which determine the specific methods of their economic forecasting are most inherent to aircraft systems, a majority of the applied questions is examined in terms of aircraft systems of various classes and purposes.

The book has been written using materials from the theoretical research by the authors and from an analysis and generalization of Soviet and foreign literature on the questions raised. The leader of the author collective is Doctor of Economic Sciences, Prof S. A. Sarkisyan. Chapter 1 was written by S. A. Sarkisyan; Points 2.1, 2.3 and 2.6 of Chapter 2 were written by Candidate of Economic Sciences, Docent E. S. Minayev; Points 2.2 and 2.4 by Candidate of Technical Sciences, Docent V. I. Kaspin; Points 2.4.5 and 2.5 jointly by V. I. Kaspin and Candidate of Economic Sciences, Docent P. L. Akopov; Chapter 3 by S. A. Sarkisyan and Doctor of Economic Sciences, Prof D. E. Starik; Chapter 4 by S. A. Sarkisyan and P. L. Akopov; Chapter 5 by D. E. Starik.

The authors would like to thank Senior Science Associate Ye. V. Tabachnaya, Engrs Yu. A. Teplov and A. S. Chernaya for the help given in preparing the manuscript for publication.

The book does not claim to be an exhaustive exposition of all the aspects of the posed problem, and for this reason the authors would be grateful for critical comments and proposals which should be sent to the following address: Izdatel'stvo Mashinostroyeniye, No 3 First Basmannyy Lane, B-78, Moscow, 107885.

CHAPTER 1: SCIENTIFIC AND TECHNICAL PROGRESS AND THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

1.1. The Scientific and Technical Revolution and Large Technical Systems

The development of technology in recent decades has shown a transition from technical devices to technical systems and this to a significant degree determines the essence of the present-day scientific and technical revolution.

The advances made in individual scientific and technical sectors, in nuclear physics and power, electronics and computers, aircraft and missile construction could be attained only by the creation of systems.

If one bears in mind modern science as a whole, in it one could scarcely find a concept capable of rivaling the word "system" in terms of breadth of use. Biologists and physicists, cyberneticians and psychologists, cosmologists and economists analyze and model a system.

The same thing can be said about modern technology. Not so long ago specialists of a corresponding specialty designed means of communications or transport and then, depending upon the specific technical parameters of this equipment, developed auxiliary facilities which would ensure their successful use. The present development stage of technology is characterized by the designing not of individual pieces of equipment but rather technical systems which incorporate all the elements essential for carrying out a certain complex function.

A modern aviation or missile complex, a production control system, a telephone network serving millions of subscribers or a large power system could be created only by considering the complex interaction of the entire system of operations and different types of equipment. All this equipment must be designed simultaneously, in a strict relationship subordinate to carrying out the basic function, and an omission in any of the elements can tell decisively on the entire system.

Large technical systems are the result of the action of fully automating the system functioning processes and the development of computers. The increased scale of activities performed by equipment, the complexity of the problems solved and at the same time the necessity of a more rapid pace of decision taking have led to a situation where the historically formed systems of control and data processing have been unable to promptly produce optimum solutions. In particular this is characteristic for systems with great operating speeds where a delay in taking an essential decision can lead to catastrophic consequences. Aircraft are a vivid example of such systems.

The complexity and diversity of the problems solved and the specific conditions of realizing them have led to a sharp rise in the number of specifications of the flight and surrounding medium which influenced the course of carrying out the set task. As a result over the last quarter of a century the number of aircraft instruments has risen by more than 10-fold while the time needed to perform aircraft control operations, due to the sharply increased flight speed, has been reduced by 6-7-fold. At the same time human response speed to equipment signals even with intense training has remained on the previous level (0.1-0.3 seconds).

As a consequence of the simultaneous execution of a range of involved functions, the necessity has arisen of automating the aircraft control process. Onboard computer equipment has appeared on the aircraft.

The process of automating the various functions performed by aircraft has developed particularly intensely in aviation where the speed factor has always played a crucial role. Here is a characteristic example from the development history of combat aircraft demonstrating the need to develop complex technical systems. The combat aircraft employed in World War II had machine gun and canon weapons and aiming was done visually. Due to the speed differences of a bomber and a fighter, the latter had an opportunity to execute several combat turns for attacking the target. However with an increase in speed the number of possible attacks declined sharply.

Analogous trends can also be traced in civil aviation. Thus, in an air traffic control system, with an increase in the amount of traffic and the number of aircraft in the air, the ensuring of flight safety by traditional methods became impossible. For example, for solving the problems involved in figuring the optimum aircraft routes for all the centrally scheduled trips observing a number of constraints such as the capacity of the routes, altitudes and so forth, it would be necessary to examine more than 10^{100} variations and choose the optimum one.

The latter can be carried out only on a basis of modern computers by creating automated flight control systems. It must be emphasized that the most important control functions in such systems and namely assessing the developing situation and decision taking, as before are carried out directly by people but the accelerated process of transforming and processing the increasing information flows as well as calculating optimum solutions using computers have become a qualitatively new phenomenon. What are the particular features which put large technical systems in a special class of systems? They are:

1) The complexity of the structure and behavior of the system, that is, the presence of such complex intertwined and overlapping relationships between the changing parameters of the system whereby a change in one variable leads to a change in many others; the presence of complex and overlapping ties between the elements in the system;

2) The irregularity of effects from the external environment and the irregularity of the very system's conduct which leads to the necessity of decision taking under conditions of ambiguity and sometimes active counteraction;

3) The presence of subsystems of an hierarchical and functional nature having their own particular goals from which the overall goal of the BTS is formed;

4) A high degree of automation making it possible by utilizing the computers to create flexible control, encompass complex dynamic processes with an enormous number of parameters and optimize the decisions being taken.

With all the diversity and complexity of the problems solved by the BTS, it is also possible to isolate features for their systematization and classification.

1. The specific features or the degree of purposefulness and specificness of the system; the degree of clarity and certainty in the formulation of the goal; the degree of its formalization, the range of goals (that is, the number and diversity of goals) and the hierarchy of goals. It is possible to isolate single-goal systems, that is, systems designed to solve one single task, multigoal systems for solving multiple tasks and functional systems which solve an individual aspect or facet of a general task.

2. The degree of integrity, that is, the degree of the permanent dependence of the component parts, elements and processes of the examined systems or stages and the directions of the tasks being carried out. Integrity is characterized by the number and diversity of harmonious links, component parts and elements of the system, by the degree of determinism in their reciprocal conduct and functioning and by other features.

3. Complexity or the degree of objective complexity; this is determined by the total number of elements and links between them, from the diversity of elements and links, from the number of hierarchical levels, from the number of functional subsystems and from other features. Depending upon the number of elements, the character of the links and the conduct it is possible to isolate the following systems:

a) Simple or small--systems with a limited number of elements (10-10^4), the links and conduct of which are a determined nature;

b) Complex or large--systems with a large number of elements (10^4-10^7) with a mass variable number of links; the behavior of such systems represents a random process which moves toward a certain limit, and for this reason such systems are of a probability sort; characteristic for them is a high degree of automation for the control processes; in particular, modern aerorocket, space-missile and other aircraft systems belong in such systems;

c) The ultracomplex or self-developing--systems with a number of elements up to 10^{30} in which successful adaptation to randomness will be carried out by the randomness of the internal structure.

4. Controllability--the degree of automation of control over the functions carried out. According to this feature it is possible to establish three basic classes of the BTS: I. Information retrieval systems (IPS). II. Automated control systems. III. Automated national-scale control systems.

Automatic telephone systems would be put in the systems of class I. Close to them in terms of the problems solved are the information retrieval systems which use an electronic computer to retrieve scientific and technical information. Here the computer is the central element of such systems and provides the link between the consumers (subscribers) and the information sources. Approximately the same principle has been used to develop and operate at present systems which locate free hotel rooms, sell air tickets (the Sirena automated ticket sales and reservation system serves 250 cities) and so forth. The data of Table 1.1 give some idea of this class of large systems in which the complexity of data transmission and processing depends essentially upon the number of subscribers (users).

Table 1.1

Subscribers	Name of System	Repository of Information Sources
Scientific workers	Retrieval of scientific and technical information	Libraries, repositories of scientific and technical papers, microfilm holdings and so forth
Visitor	Hotel reservation	Hotels
Leadership, administration of sector or ministry	Obtaining information on course of carrying out production plans, material-technical supply and so forth	Enterprises under the ministry
Aeroflot reservation clerk	Air reservation	Aeroflot ticket service which has tickets and monthly flight schedules

A common feature of the class II systems is traffic control where a person acts as an operator controlling a process, as an inspector or a diagnostician in a closed control loop both for the entire system and its individual subsystems. Among such systems one could put:

a) An automated air traffic, take-off and landing control system which ensures safety in carrying out scheduled, training and other flights within the limits of a given airport and simultaneously maximum utilization of the runway capacity;

b) The control system for a large aircraft or space device;

c) An automated control system for production processes (production processes in petrochemistry, the cement industry, the extraction of metals from ores, the rolling of metals and so forth;

d) The control system for energy or transport systems and so forth.

A basic feature of the class III systems is the use of class-II systems which have varying functional purpose within a single system unified by a common goal. Depending upn the degree of detailing for the component elements and functions of the class II systems, by a class III system one can understand, for example, both the entire automated air traffic control system encompassing the territory of an entire country as well as an individual subsystem concerned solely with the questions of dependable communications between the airfields comprising the air traffic service. The following basic features of the class III systems can be established:

a) Control of the processes is significantly more complicated than in the class II systems; entire complexes and associations of the class I and II systems which carry out different functional tasks can serve as objects of control;

b) The hierarchy of the system's structure and the higher the level of hierarchy the less contact with the specific functions performed by the lower rank systems;

c) The presence of class I information retrieval systems as data sources;

d) Control of the major operations using various automatic data processing and display devices.

Thus, the basic feature of complex systems is information processes linking the individual elements into a single whole for ensuring optimum control. For precisely this reason a system is not a simple combination of its own subsystems but rather possesses particular properties which none of its individual parts has.

Cybernetics, information theory and algorithm theory are concerned with the questions of controlling the BTS. However, in the process of developing the BTS, a multiplicity of important problems arises going beyond cybernetics and the other above-mentioned sciences. One of them consists in creating an economically optimum system in terms of its set functions. The development of science and technology provides an opportunity to create a great diversity of technical devices or elements of the BTS capable of carrying out qualitatively uniform functions. Due to the differences in the physical processes which ensure the realization of a certain function, these devices possess different functional characteristics and a design or technological appearance, they consume different types of energy and so forth. The listed features determine, on the one hand, the operational efficiency of the technical devices and, on the other, the cost of their creation, production and operation.

The diversity of functionally equivalent elements for the BTS gives rise to an even greater diversity in the variations of constructing it as each of these is capable of realizing the set behavior. The variations of the BTS synthesized in a certain initial range of devices or elements which are indistinguishable in terms of functional features will possess their own characteristics of cost and effectiveness. The latter gives rise to the problem of selecting a preferential alternative out of the multiplicity of systems which realize the set behavior. The choice of alternatives can be made objectively only under the condition that this is done on the basis of analyzing the economic consequences of developing the BTS and the entire spectrum of expenditures on their creation, production and operation.

The problem of selecting an economically optimum system is solved by the methods of general systems theory and systems analysis, the theory of the economic effectiveness of capital investments and new technology and scientific and technical forecasting.

1.2. General Principles of Research and Analysis of the BTS

The principles for the research and analysis of large technical systems are based upon the methodology of studying complex objects and this comprises the foundation of so-called systems theory. General systems theory as a concept was proposed in the 1930's by the Austrian biologist L. von Bertalanffy at a philosophical seminar at the University of Chicago. He developed an "organismic" viewpoint in proposing to view living organisms not as aggregates of cells which, in turn, consisted of colloids and organic molecules but as an organized unified system.

An examination of objects representing a certain aggregate of interrelated and interdependent elements as a unified (unifiable) single whole for carrying out a complex function has been termed a systems approach. The systems approach gained universal recognition and was fruitfully employed in a study and analysis of various material objects and processes of their development due to its dialectical essence. This approach represents the extension of well-known methods of Marxist dialectics which view nature as a single related whole in which phenomena are interdetermined and interdependent to economic, biological, cybernetic, technical and other systems which are component parts of the material world.

The methodology of the systems approach is the unifying principle which makes it possible to extend its principles to diverse scientific areas. The methodology is based upon the principles of the integrity of the studied object and the principle of isomorphism. The principles of integrity are related to the hierarchy of the system's structure and make it possible to represent the complexity of the studied system in a broken down form. The principle of isomorphism[1] is used for an analysis of the laws which explain the inner similarity of objects and structures of different nature and purpose.

In the sphere of technology, a systems approach has been effective due to the fact that modern technology, as has already been pointed out, in its essence is systems technology as:

a) Its basic object is various types of systems (aviation, missile, space, production control systems, communications systems and so forth);

b) The process of creating technical systems is itself a system in which is carried out the coordinated work of numerous prototype design, scientific research, production and operating organizations which, in turn, represent independent systems;

c) The process of producing the technical systems or their construction is carried out in a certain system and, as a rule, this process involves numerous enterprises which are elements of a complex production system;

---

[1] By isomorphism one understands a uniform similarity between objects from the viewpoint of their structure, the relationships between the structural elements as well as between the objects and the external environment.

d) The process of the functioning of technical systems is a system in which there is an interaction not only between the elements of the technical system but also between the different specialized systems.

Since the main distinguishing feature of a large system is the close link of all its elements and parts, a systems approach to the analysis of a BTS means a consideration of these relationships, a study of the individual objects as structural parts of more complex systems and the ascertaining of the role of each of them in the overall process of the functioning of the BTS.

Let us examine certain concepts of systems theory which are most often used in the text below. For this let us formulate again the concept of a "system" and isolate its basic properties which distinguish a system from any other aggregate of elements.

In general systems theory, by a system one understands an aggregate of objects or elements which possess certain properties and are interconnected and by these interconnections the system is unified into a single whole. The system possesses a certain structure which allows a breaking down of the hierarchy of elements. It interacts with an external environment and can be viewed as an element of a broader system that is superior to it. The structure of a system is such that its elements possess the properties of a subsystem in relation to it. The system is designed to perform a certain activity which can be broken up into a number of interrelated operations. From the definition of a system it follows that the most important concepts in general systems theory are the element, operation, external environment, structure and hierarchy.

An element of a system is what lies at the basis of the hierarchy in breaking down the system and cannot be broken down further.

In accord with the role that the system's elements play in the process of achieving the set result, the so-called system central element[2] is isolated among them (the elements) and by this one understands the entity (the aggregate of interrelated elements) capable of performing an elementary operation.

An operation is an aggregate of actions aimed at achieving a certain goal. In the process of performing an operation, the central element will be linked to other parts of the system and the interaction with them carries out the operation. However, the characteristics of the central element have a determining impact on the functional properties of the entire system.

Any system operates in a certain environment. The environment is the aggregate of all elements where a change in the properties of these influences the system as well as those objects the properties of which are altered as a result of the system's

---

[2] In certain instances, the term "central subsystem" is used. For example, the aircraft is the central element (subsystem) of a system of aircraft which would include the aircraft and the ground facilities such as the airfield, controls, communications and so forth.

behavior. For this reason, both the system as a whole as well as each element in the system have inputs which characterize the actions of the environment on the system and its elements and outputs which characterize their effects on the environment.

The interaction of the system with the environment as well as that of the system's elements with one another can be represented by structural models or functional models.

A structural model, depending upon the aim of analyzing the system, can be of three types: an external model in which the system is represented in a canonical form and all its links with the environment are expressed by inputs and outputs; a hierarchical model in which the system is broken down by levels according to the principle of the subordination of inferior levels to superior ones; an internal model in which the composition and relationship between the system's elements are shown.

The functioning of the system can be represented by the following: by a model of the system's life cycle characterizing the processes involved in the system's existence from the genesis of the idea of its creation to its "death" (the ceasing of functioning); by a model of the system's operation representing the aggregate of processes involved in the system's functioning for its basic purpose.

All these models characterize the system's method of action (the method of existence and functioning) in space and time.

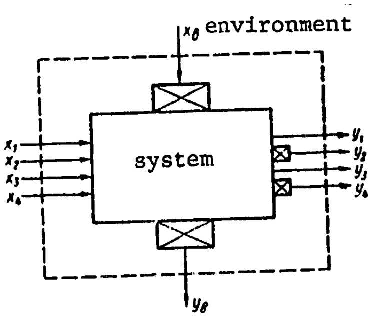


Fig. 1.1. Canonical model of a technical system with inputs (outputs): $x_1(y_1)$--information; $x_2(y_2)$--energy; $x_3(y_3)$--object; $x_4(y_4)$--personnel; $x_B(y_B)$--disturbances; ⊠--filter.

Let us examine the particular features of canonical models of systems. As is shown in Fig. 1.1, the basic input vector components are: x_1--the information input which controls the activity of the subsystem or is subject to processing by the system; x_2--the energy input which ensures the development of the system or maintains it at the set level of productivity; x_3--the material or object input which represents the flow of materiel to be processed by the system (the material means for the operations performed by the system); x_4--the personnel input which provides the system with personnel prepared for participating in the functioning processes.

The designated inputs represent organized inputs and their presence is ensured by the purposeful activities of people. In addition to the organized inputs there are also unorganized ones which, as a rule, impede the system's activities or these might be called the disturbance inputs x_B coming from the environment (interference, noise, constraints and so forth). Thus, the input of a BTS is a vector

$$\overline{x} = (x_1, x_2, x_3, x_4, x_B).$$

Each input will have several components so

$$x_i = (x_{ij}), \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m;$$

$$x_{ij} = (x_{ijg}), \quad g = 1, 2, \ldots, p,$$

where i--the type of input; j--nomenclature of input; g--source of input.

The result of the system's activities, the output vector $\overline{y}$, can be characterized by analogous components:

$$\overline{y} = (y_1, y_2, y_3, y_4, y_B),$$

where y_1--the information output characterizing the result of the system's information activity;
y_2--the energy output characterizing the transfer of energy from the system to the environment and the loss of the system's elements (the exhaustion of their life, the nonconformity to demands or flaws) as well as production wastes;
y_3--the object output characterizing the result of the purposeful action of the system (what the system produced);
y_4--the personnel output characterizing the movement of personnel;
y_B--the output disturbance characterizing the system's ancillary actions on the environment.

Obviously, as is the case for the inputs, the output vector components can be represented in the form

$$y_i = (y_{ij}), \quad i = 1, 2, \ldots, k; \quad j = 1, 2, \ldots, \ell;$$

$$y_{ij} = (y_{ijg}), \quad g = 1, 2, \ldots, s,$$

where i--type of output; j--nomenclature of output; g--purpose of output.

The characteristic inputs and outputs of a passenger airplane as a system are shown in Table 1.2. An analogous approach to systems analysis using canonical models can also be applied to production systems. The characteristic inputs and outputs of a system in terms of the production of large technical systems are shown in Table 1.3. From Tables 1.2 and 1.3 it follows that, regardless of the difference in the structure and the functions of the designated systems, their inputs and outputs keep fully within the given input and output classification. The latter, in particular, means that the compared systems, regardless of their differing nature, are isomorphic from the viewpoint of the external structure.

A study of a canonical model in terms of a specific system makes it possible to disclose the relationships of the systems. The inputs and outputs here can be expressed by parameters which comprise the system's functional model (the model of the operation).

Table 1.2

Inputs or Outputs	Name of Passenger Aircraft Inputs and Outputs				
	Object x_3, y_3	Energy x_2, y_2	Information x_1, y_1	Personnel x_4, y_4	Disturbances
Input	Passengers and cargo to be carried	Fuel (ground input) and oxidant (in-flight input), spare parts for replacing worn ones (ground input)	Flight assignment, take-off commands; in-flight information from ground control and navigation systems; landing commands; requests and so forth	Flight crew, ground service personnel	Bad weather conditions holding up take-off; appearance of malfunction in system; storm formations enroute
Output	Passengers and cargo delivered to destination	Energy losses in atmosphere; loss of system's elements as used up or due to malfunction	Information on results of trip; on state of system; on the flight; the carrying out of the mission; replies to requests and so forth	Temporary departure of flight crew from system (for rest) Departure of crew members necessitating change	Shock wave in crossing the speed of sound (effect on population); jettisoning fuel in emergency

Table 1.3

Inputs or Outputs	Name of Inputs and Outputs for Producing BTS				
	Material	Energy	Information	Personnel	Disturbance
Inputs	Raw products, materials and semi-finished products; parts, assemblies and units manufactured by subcontracting	Production equipment, buildings, installations to replace worn out ones Electric power and heat necessary for production operations; fuel, various fluids and gases for flight testing	Plan specifications for manufacturing aircraft; subcontracting plans; technical and engineering specifications; information on the formation of inputs; information on the results of operating the produced aircraft	Personnel trained for work at enterprise; personnel requiring retraining and training by enterprise	Unforeseen changes in technical specifications for manufacturing the articles; meteorological conditions impeding testing; and so forth
Outputs	Finished aircraft; spare parts; articles manufactured by subcontracting plans	Loss of production equipment and fittings; loss of buildings and installations; production wastes	Technical descriptions, instructions; product certificate; report documents; information from operating organization	Personnel under planned transfer; personnel losses	Output of defective products; violation of subcontracting delivery system

In the process of systems analysis and forecasting it is important to know now only their link with the environment but also their structure. In considering that the different natured systems are isomorphic in terms of their inner structure, let us examine the principle of constructing hierarchical models for the inner structure of systems using the example of aircraft systems.

In the general instance an aircraft system is divided into subsystems of a certain rank (Fig. 1.2, a). In the diagram S_0 is a certain supersystem (for example, the transport system) in which the aircraft system is included as a subsystem (a system of the first rank).

The system of aircraft S_2 consists of several aircraft subsystems $S_{21}, \ldots, S_{2m}$ and several support subsystems for their functioning $S_{2(m+1)}, \ldots, S_{2(m+n)}$, which, in turn, will be the subsystems of the second rank. For example, in an air transport system the aircraft systems of the second rank can be the air transport systems in the economic regions of the nation while the support systems for their functioning are the air traffic control system, the system for the development and overhaul of aircraft, the material and technical supply system and so forth.

The second-rank aircraft systems, in turn, can be broken down into the third-rank systems which also will make up several aircraft subsystems and several intermediate support subsystems for their functioning. Here the latter will not be identical with the support systems of a higher rank.

Finally, each third-rank aircraft system can be broken down into several aircraft systems and support systems of the fourth rank. Such a breaking down can lead to ensuring the identicalness of the inferior rank systems.

If the systems $S_{2mi1}, \ldots, S_{2mij}$ are identical in terms of the composition of the aircraft and their functions, then the inner structure of each of these identical systems can be represented by a single scheme (Fig. 1.2, b). An identical aircraft system includes the aircraft (aircrafts), the take-off system (airfield), the control system and the repair and support system. The aircraft can be divided into expendable and reusable subsystems and so forth.

An analysis and assessment of the systems for the purposes of forecasting their development are the basis for the scientific choice and disclosure of the relationship between the goals of the system, the means of achieving them and the resources. The basic goal of systems analysis is the taking of a decision on the ways to improve the system or process. A decision describes the difference between two states and determines the method for moving the system to a new state. The implementation of the decision is the process of moving the system to a new state.

In terms of the contents of the analysis problem, systems can be divided into four types: the problems of optimizing the designed parameters of the system; the problems of selecting a preferential alternative (the selection of a preferential system); the problems of allocating the assigned resources in the stage of making up the complex systems (in forming a "mix") under the conditions of an ambiguous situation; the problems of allocating the resources available to the systems (for achieving the goals of the operation in a specific situation). The first three problems are the problems of developing the systems and the last is the problem of using them.

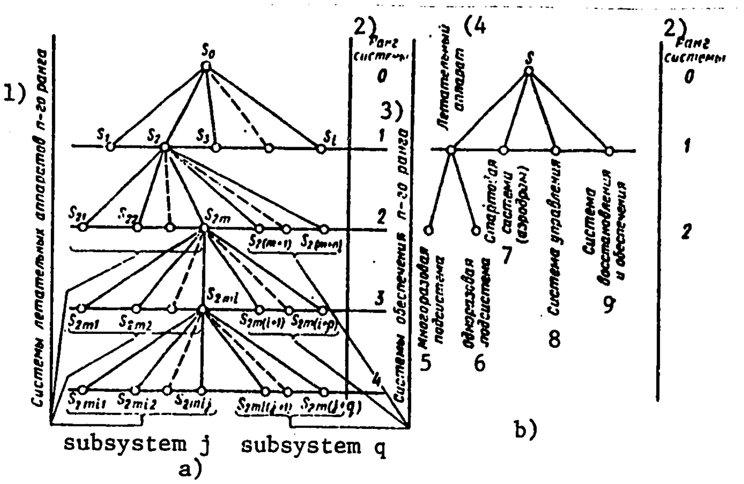


Fig. 1.2. Tree of system's hierarchical structure for aircraft

Key: 1--Aircraft systems of rank n; 2--System rank; 3--Support system of rank n; 4--Aircraft; 5--Reusable subsystem; 6--Disposable subsystem; 7--Take-off system (airfield); 8--Control system; 9--Repair and support system

The process of analyzing large technical systems includes the following areas of research.

1. Determining the ultimate goals of the system.

2. The working out of alternative methods and means for achieving the set goals and variations of systems from which the most preferential must be selected.

3. Ascertaining the required resources to implement the designated alternatives and the constraints on them.

4. An analysis of the interaction of the goals, alternatives and resources, including interrelated events such as: the selection and formation of the evaluation criterion and the constraints which define the area of possible decisions; a comparison of alternatives by a criterion, including an optimization of the decision with an analytical form of a criterion; defining the ambiguities and an analysis of their influence on the calculation results; judgments complimenting the analytical analysis; taking a decision on the choice of the preferential variation of the system considering additional information on possible situations, interacting systems, available resources and so forth. If the results are unsatisfactory, then a decision is made to carry out a new cycle of analysis with a revision of the set goals and the elaboration of new alternatives and resource constraints.

The obtained decisions are the basis for elaborating the specific, operational program and economic forecasts. The integrity of the compiled forecasts will be largely determined by the objectivity of the criteria used as the basis for selecting the systems and the considered constraints, by the correctness of the formal methods, by the depth of analysis and so forth. However, along with this of importtant significance will be how complete and thorough analysis is given to the patterns inherent to the development processes of the BTS and the relationship of these processes to the overall development trends of science and technology in related spheres of activity employing different systems as well as in the sectors of material production.

1.3. Particular Features in the Development of Large Technical Systems

Large technical systems are developing systems. In studying the BTS it is essential to bear in mind two aspects of systems development: genetic, that is, the study of a system in its development, and functional or the study of the actual actions of a system and its functioning. From the viewpoint of the methodology in economic forecasting of interest is the genetic analysis, that is, the examination of the origin and particular features of a system's development.

There are two approaches to explaining the nature of the processes in scientific and technical development: ontological and teleological. The sense of the former approach is that the development processes are viewed as a manifestation of a self-developing synamic process or the result of activities by a self-developing system. In other words, scientific and technical progress is viewed as a response to the opportunities and problems confronting science and technology. Here is assumed the presence of factors which are internally inherent to science and technology and cause the process of scientific and technical development.

The supporters of this view refer to the fact that the inventions which have caused major consequences are accidental and not determined by external causes, or, in any event, are determined by certain concealed factors which are outside the sphere of action of the main driving forces of history (the discovery of the antibiotic properties of penicillin, the discovery of radioactive decay, the invention of the laser and so forth).

The teleological viewpoint holds that scientific and technical progress is considered as the result of an objective process determined by social need or a great economic demand. The primacy of the external (social) effects on scientific and technical progress assumes that the rate and direction of the latter can be predicted only to the degree that scientific and technical progress itself is the consequence (that is, the reaction) of changing needs or demands externally superimposed on the system of research and development. In other words, if a social need is recognized, then the technical means for satisfying it can be provided.

These two approaches are diametrically opposite viewpoints and while the former fully excludes the possibility of controlling the development processes the latter assumes that these processes are fully controllable.

How do large technical systems develop? From the viewpoint of establishing the patterns in the scientific and technical development of the systems of greatest interest is an examination of the class of competing BTS, the examples of which would be aircraft systems.

The schemes for the functioning of a system is represented by its canonical model which depicts the aggregate of factors characterizing the process of its functioning through an external structure, the system's inputs and outputs. The latter are determined by the relationships between the designated system and the environment.

Characteristic of technical systems are three groups of inputs and outputs, the information, energy and material. Their content depends upon the presence of a competing system and the relations between competing systems.

Characteristic of the relationship between competing aircraft systems are two periods of life: the period of their nonconflicting competition and the conflict period. The systems will have different inputs and outputs in accord with these periods.

The canonical model of a competing system can be most fully represented for the conflict period. In this instance all the inputs can be divided into three groups (Table 1.4): those depending upon the researcher X_1, X_2, X_3, X_4, those depending upon the competing system X_5, X_6 and those depending upon nature (if nature does not operate as a competing system) X_7. The results of the transformation of the inputs by the system will describe the system's outputs.

Let us describe in somewhat greater detail the significance of the inputs and outputs for a certain aircraft system S.

For each such system the basic object of effect is a certain aggregate of goals (the system goal) which will be the basic content of input X_5. The results of the system's effect on the system goal will be described by the output Y_5. In turn, the effect of the system goal on the designated system will be described by the input X_6 and the change in the state of the latter as a result of this effect by the output Y_6.

Thus, the efficiency level of a competing system depends, on the one hand, on the conformity of the controllable inputs X_1 and X_2 to the needs of the system, and on the other, upon the state of the inputs regulated by the competing system, X_5 and X_6. For this reason the development of aircraft systems has a competitive nature. Each of the competing sides endeavors to increase the efficiency level of its system and thereby reduce the efficiency of the competing side's system. Under these conditions, even during the nonconflicting period, relative efficiency of the competing system shows a wave-like nature. After one of the sides has improved its system for the purpose of raising its efficiency, the opposite side endeavors either to minimize the gain in efficiency achieved by the competitor by countermeasures or to make its system as advanced as the competitor.

Consequently, it can be stated that the development of competing systems has a dual nature. On the one hand, this development is a response to a change in the state of the system goal, that is, the competing system in order to prevent a decline in the efficiency level of one's system by employing countermeasures. At the same time, a change in the efficiency level of the competing system can occur and often does occur as a result of spontaneous discoveries and inventions (for example, the development of more advanced equipment, semiconductor electronics and so forth). Thus, the motivating force in the development of a BTS is simultaneously the social needs and the inner possibilities of scientific and technical progress which open up new, previously unknown prospects for their improvement and application.

Table 1.4

Inputs	Outputs
X_1--information, determining program of system's work	Y_1--information, describing results of system's information activities
X_2--energy, (resources ensuring development, safekeeping, functioning and repair of system)	Y_2--energy not depending upon competing system (loss of system's elements, consumption of resources as a result of system's functioning)
X_3--conditions and constraints, imposed by interacting systems	Y_3--conditions and constraints, imposed by the results of system's functioning on interacting systems
X_4--conditions and constraints, imposed by national economic interests	Y_4--effect of system on national economy
X_5--object, or the system goal (the objects of the system's effect)	Y_5--goal or specific (the results of the system's functioning for its basic purpose and the change in the state of the system goal)
X_6--competing,	Y_6--energy dependent upon competing system (change in system's state, loss of system's elements as a result of counteraction and resources on repair of system)
X_7--conditions and constraints, imposed on system by nature	Y_7--effect of system on nature

What has been said predetermines the strategy of analysis and building of systems whereby the choice of the optimum directions of systems development is carried out proceeding from the set goals of their functioning but considering the means (possibilities) which are provided by scientific and technical progress.

The development of new technology can have an abrupt or evolutionary nature. From this viewpoint scientific and technical progress consists of definite stages (markers) which differ qualitatively from one another. These stages are not absolute and their relativity consists primarily in the fact that each new stage is a dialectical negation which includes an aspect of succession, maturation and development and the synthesizing of certain elements from pervious stages [28]. The transition to a new stage is not a single act or a boundary point of development as the technical and scientific revolutions or their stages can be superimposed one on the other.

The relativity of the stages and the revolutionary periods of science and technology, the links between them and their possible superimposition--all of this does not show that technical and scientific progress is a continuous chain of revolutionary changes. The pace of scientific and technical progress periodically alters. Scientific and technical development always includes not only the abrupt shifts and revolutions but also periods of evolutionary movement. Abrupt development occurs in the transition to qualitatively new physical phenomena and materials and is expressed in the appearance of new classes of systems.

The change from cycles of evolutionary technical development to abrupt shifts in its functional properties can be easily traced from the example of the change in the speed of means of transport. The functional parameters evolve within the limits of a certain class of systems which are unified by common principles in the functioning of the major subsystems [airplanes with piston engines (PD), aircraft with gas turbine engines (GDT) and missiles].

In the general case, the evolutionary process of functional characteristics in systems undergoes a number of sequential phases: the phase of embodiment, the initial phase, the phase of intensive development (maturity) and the phase of obsolescence (Fig. 1.3, curve 1).

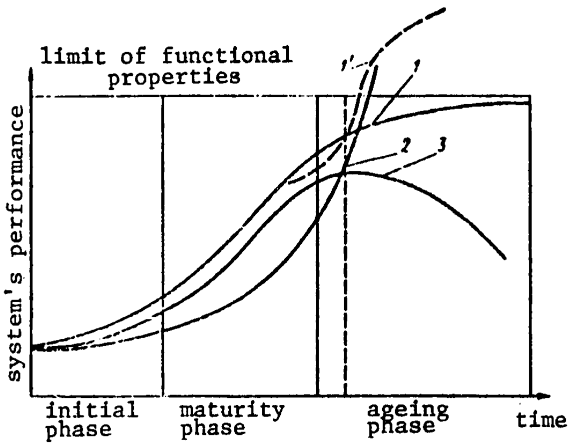


Fig. 1.3. Dynamics of most important BTS performance:

1--Functional properties; 2--Cost; 3--Efficiency of system

The embodiment phase which precedes the appearance of the prototype includes research on the physicochemical principles of the system's functioning, the methods of creating a useful effect based on the results of the theoretical and experimental research and the possible spheres of the new system's application.

The initial phase, or as it can be called the incubation period, coincides in time with the beginning of materializing the scientific and technical ideas. During this stage the first models appear of the functional subsystems which employ new physical and physicochemical processes which fundamentally distinguish these subsystems from their predecessors.

During the incubation period the basic efforts of the research and development organizations are aimed at ensuring the stability of the occurring processes as well as studying new phenomena which appear in testing the BTS which employ such subsystems. As a rule, the prototypes of the new BTS do not go into industrial production or operation. More advanced articles which employ the same functioning principles are developed on their basis and undergo experimental testing. The growth rate of the functional parameters during this period is still slight but continuously increases.

The intensive development phase can be termed the period of the system's maturity. The maturity period encompasses the time from the appearance of the first industrial models to the moment when the potential provided by the nature of the occurring processes is virtually exhausted. Characteristic of this period is the highest growth rate of the functional parameters. It is essential to note in particular that a series of modifications and modernizations occur in this period. The functional parameters increase during this period by a slight amount in comparison with the base article but here the spheres of use of the BTS are substantially widened.

Over time the increase rate of the parameters gradually declines and the moment comes beyond which the increase rate in the parameters begins to drop continuously. This is caused by the influence of impeding factors for the given type of equipment (for example, the piston engine restricted the possibility of developing supersonic aircraft).

The last phase in the development of the BTS is the equipment obsolescence phase when the possibilities of further improving the equipment are exhausted in terms of the old fundamental bases and the growth rate of the functional possibilities decline sharply. During this period, as a rule, there begins the materialization of new scientific and technical ideas aimed at broadening the theoretical limits of functional characteristics which restrict a further rise in operating efficiency and a broadening of the sphere of use of the BTS. The latter is accompanied by a qualitative shift in the functional performance of the BTS subsystems and properties (Fig. 1.3, curve 1').

A combined examination of the change patterns in the functional parameters of technical systems and their cost estimates[3] makes it possible to spot the most important feature in the system's changed efficiency which can have a decisive impact on its development.

Numerous research has shown that the cost estimates of systems respond regularly to a change in their functional properties and parameters (this question will be examined in detail in Chapter 4). The evolution of functional properties, in being accompanied, as we have seen, by a greater complexity of the systems, leads to an intensive rise in their cost estimates (Fig. 1.3, curve 2). Here, while the growth of the functional properties is restricted by the system's nature and as a consequence of which its development rate moves to zero, the system's cost rises exponentially.

---

[3] By cost estimates here and below we understand expenditures on development (research and development), industrial production and operation of the systems.

The efficiency of a system, as a result of the interaction of the functional properties and cost of the system, with the growth of the functional parameters initially increases and then, upon reaching a certain maximum value, begins to diminish sharply (Fig. 1.3, curve 3). Consequently, it is possible to speak about a certain area of an economically optimum existence of the system beyond which it is necessary to use fundamentally new systems for the same purposes. The locating of these areas is one of the most important tasks in economic forecasting as it opens up an opportunity to effectively control the scientific and technical development processes of the BTS.

The development processes of a certain type of modern technical systems cannot be viewed in isolation from other systems as well as outside of the scientific and technical development processes in the sectors involved in their creation and production.

Thus, in selecting the development strategies for a certain class of systems it is essential to consider not only the direct result of this development in the form of the greater efficiency of the system. It is also important to take into account the side effect which can appear as a consequence of implementing the results of the given system's development in systems of a different class or purpose.

Scientific and technical progress is expressed not only in a change in the properties of the BTS and in the use of the results from this development in other areas of human activity. The ensuring of the set functional properties of the systems often requires the employment of fundamentally new means and methods of their creation and production. This leads to a situation where in the process of scientific and technical development the material and technical base of the sectors producing the new systems undergoes profound changes. Fundamentally new equipment and production processes are introduced and these provide high precision in the working of the parts and joints as well as high purity and uniformity of structures both in working traditional and fundamentally new materials.

In parallel with this an improvement occurs in the processes of creating systems for the purpose of raising labor productivity, reducing labor intensiveness, shortening the cycles and increasing the efficiency of control. The enterprises which produce the BTS elements automate the processes involved in controlling conditions in the heat-treating and plating shops, the processes of milling part contours using hydraulic and electric tracking systems and machine tools with program control are evermore widely used. Machining is replaced by cold upsetting, cold extrusion, electroupsetting and rotary working. Ultrasound and photoelectronic, magnetic powder and capillary methods are employed for quality control of the initial materials, castings, forged pieces, finished articles, joints and assemblies with a high degree of precision and reliability. The organizational management structures are being improved and automated production control systems (ASUP) and automated development control systems (ASUR) are being introduced.

Thus, the process of developing the BTS is a multi-aspect one involving a rise in the technical level in the sectors involved in their creation and production as well as in the related sectors. All of this shows the necessity of viewing the development processes as a whole and considering the relationship and intercausality of the development rates in the individual areas. In other words, a systems approach should be applied to analyzing the BTS development processes.

1.4. The Life Cycle and Scheme of the Analysis Process of a BTS

The large technical systems exist in space and in time. The time period from the plan to create a system until it is taken out of operation is termed the life cycle of a system. It includes several stages, each of which consists of a number of events and levels. The duration of a system's life cycle depends upon its purpose and technical potential. The basic stages of a BTS life cycle (Fig. 1.4) are: scientific research; prototype design work; series production; operation. The beginning of a system's life cycle is preceded by the phase of social, economic and scientific-technical forecasting. This includes the range of work on shaping the tasks of the systems and assessing the possibilities of science and technology over the long run.

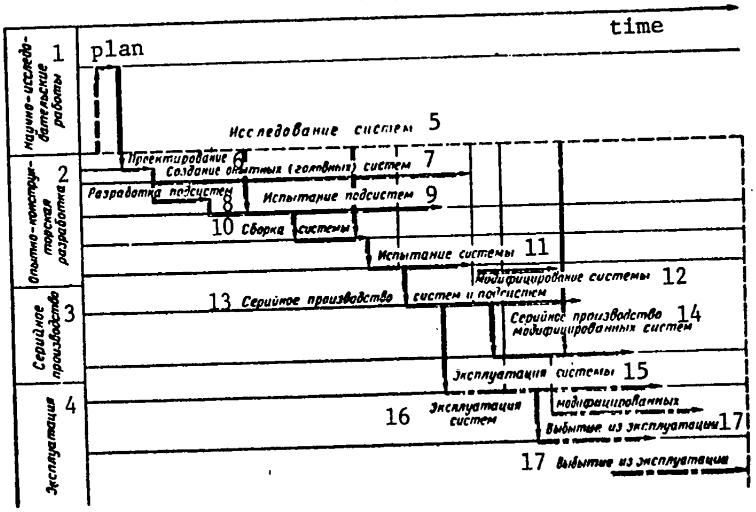


Fig. 1.4. A System's Life Cycle

Key: 1--Scientific research; 2--Prototype design work; 3--Series production; 4--Operation; 5--Systems research; 6--Designing; 7--Creation of prototype (head) systems; 8--Elaboration of subsystems; 9--Testing of subsystems; 10--Assembly of system; 11--Testing of system; 12--Modification of system; 13--Series production of systems and subsystem; 14--Series production of modified system; 15--Operation of system; 16--Operation of modified systems; 17--Taking out of operation.

Scientific research starts with the plan of the BTS (the phase of shaping the concept). The genesis of the plan starts with an awareness on the part of the organization in charge of utilizing the system for its basic purpose of a need to develop or replace the existing systems because of a widening or change in the nature of the tasks or the development of a fundamentally new system caused by the appearance of new tasks. Thus, an awareness of the new tasks and new conditions is the starting point of the plan.

The initial prerequisites for the genesis of the plan are fundamental changes in the nature of operations which shape the basic principles (the doctrine) in the sphere of the systems' functioning. Here the doctrine operates as the organizing principal. In turn, the successes in developing the new systems and the appearance of fundamentally new types of systems definitely influence the content of the doctrine. Thus, in the process of creating the systems there is a constant interaction between theory and practice, as follows: new tasks--doctrine--the plan for the system--new system--doctrine.

The forming of the plan includes a series of events the basic ones being: research carried out by the client for the purposes of analyzing the new tasks and elucidating the demands to be made on the systems designed to solve them; the shaping of the initial tactical-technical requirements (TTT) for the new systems in considering the nature of the new tasks and the scientific and technical possibilities forecasted for the immediate period (it is important that the demands reflect as fully as possible the goals which the new system seeks to attain and provide the designers and researchers with room for searching for rational ways to solve the new problems); research conducted by scientific and industrial organizations in the aim of seeking out new scientific and technical principles and ways for solving new problems; the elaboration of several variations for the initial design of the system, that is the preliminary project (predesign project) for the purpose of formulating the system's appearance, the basic relationships, the ways for solving the basic technical problems and the required resources for the creation and functioning of such a system; research on the efficiency and optimization of its parameters for the purposes of choosing the preferred variation.

The end result of the plan stage is proposals or recommendations on solving the problem and these would include the content of the plan in the form of the description of the system, the volume and sources of resources required for its creation and functioning and an estimate of the development and production times.

For choosing an optimum system it is essential to work out not only several alternative systems within one plan (several alternative subsystems within one system) but also several alternative plans. The alternative plans would include fundamentally different systems the commonness of which consists only in the commonness of the pursued goals.

The second stage in the life cycle of a BTS is the prototype design work (OKR) which includes the designing, manufacturing of prototypes (prototype production) and the testing of systems. As a rule, by the start of designing the less preferential versions of the systems have been weeded out and designing is carried out with a smaller number of variations.

The system's designing starts with an adjustment of the tactical and technical requirements made on the system. In working out the projects for several variations of systems there is an alternating of the process which follows the scheme of synthesis--analysis--systhesis--analysis and the discovery of new possibilities is not to be excluded. For this reason the client's requirements here must be considered as a guide for the areas of the search although they basically should already govern the developers. It must be pointed out that designing also presupposes the continuation of research on new problems discovered in the process of drawing up the plan and in designing.

Designing ends with the elaboration of the working drawings. The system's analysis carried out at this stage has specific features. In the first place, the analysis of the system in the designing is carried out on a theoretical basis without testing it out in a full-scale experiment. The testing provides an opportunity to check out the conformity of a whole series of calculated initial data and conclusions to the experiment. Substantiation of the conformity of the individual calculated parameters to the experiment provides great certainty of the analysis' correctness.

Secondly, in selecting the preferred system in the research and design stage it is very difficult to assess the expenditure of resources on the prototype and series production and operation of the systems with sufficient precision and reliability. This can lead to the taking of incorrect decisions. This can be done with much greater accuracy and reliability from the results of the actual expenditure of resources on the development of the prototype system. For this reason at the given stage a specific analysis of the system should be run in taking the decision about the series production program.

In the stages of the series production and operation of the systems there will be: the production of the subsystems, the assembly and installation of the systems as a whole, the functioning of the systems and the maintaining of them in a state of technical working order and functional readiness as well as the repair of the systems. The operation of the systems makes it possible to finally assess the theoretical research carried out in the process of creating the system as well as to improve the algorithm and methods of system analysis. The life cycle of a system ends with its taking out of operation as a consequence of obsolescence. A system, as a rule, is modernized by replacing some of its elements and by developing others.

As can be seen from the description of the basic stages in a system's life cycle, the analysis and assessment of systems are carried out in all stages starting with the formation of the plan and ending with the decision to take it out of operation.

The analysis and assessment of aircraft systems in the interests of forecasting their development in the early research and design stages are carried out under the conditions of an ambiguity of the situation and initial data and the presence of resource constraints. These conditions in the selection of the technology have led to the rise of a new scientific discipline, systems analysis, as a methodology for selecting systems under conditions of ambiguity and resource constraints. Systems analysis, in

Systems analysis, in being based on systems theory and using the mathematics of operations research, compliments them in its logical methods of decision preparation under the conditions of the ambiguities developed by decision theory. Systems analysis is the basis for a scientific choice and the elucidation of relationships between goals, the means for achieving them and the resources.

In comparison with operations research which provides a quantitative assessment of the results of systems use in a specific operation by using strict mathematical methods, systems analysis recognizes such an assessment as insufficient for selecting the preferred variation as a result of the presence of a number of ambiguities. For this reason the solution to the selction problem is supplemented by other methods, namely: by judgment methods based on logic and on formal experience and by engineering methods. In the latter the prime role is assigned to the art of recognizing common interrelated development patterns of systems and situations.

The process of systems analysis includes the following research areas:

1) Determining the ultimate goals of the system;

2) The elaboration of alternative methods and means for achieving the set goals and the variations of systems from among which the most preferential must be selected;

3) The elucidation of the required resources for implementing the designated alternatives and constraints in them;

4) An analysis of the interaction of the goals, alternatives and resources including interrelated events, such as: the choice and shaping of the evaluation criterion and the constraints which define the area of acceptable decisions; a comparison of alternatives using the criterion, including the optimization of the decision with an analytical form of the criterion; elucidation of the ambituities and an analysis of their influence on the calculation results; judgments or logical analysis complimenting the analytical analysis; the taking of a decision on selecting the preferred variation of the system considering the additional information on possible situations, interacting systems, available resources and so forth; if the results of the analysis are unsatisfactory, then the decision is taken to carry out a new cycle of analysis with a revision of the set goals and the elaboration of new alternatives and resource constraints.

The multiplicity of states in which a system is found during its life cycle also determines the necessity of a continuous systems analysis process. As a result of the increase in the amount of information and the degree of its reliability, one can speak of a multistep (iterative) process of systems assessment. As a first step one might point to preliminary analysis based on judgments and simple analytical models in the course of which the required information on the possible goals and areas for searching for alternatives, on operations models and so forth will be more fully disclosed.

The basic components in the systems analysis process are: the goal, operational and design research, an analysis and selection of the criterion and the constraints, modeling of resources, the criterion function and constraints and the selection of alternatives or the optimization of the system.

The goal-oriented research consists in elaborating the alternative goals and choosing the preferential alternative. The selection of the tasks (goals) the fulfillment of which should be ensured by the system is either the result of a systematic analysis of the dynamics of the tasks which arise as the situations change or the result of the generalizing of the experience and views existing on a superior management level.

The process of defining the goals is subordinate to certain rules of which we would point to the two most important. In the first place, the tasks of the interior-level systems should be compatible with the tasks of the superior level systems and, conversely, the tasks of the superior-level systems should be synthesized from the tasks of the inferior-level ones and stem not only from the needs but also from the possibilities of the systems. Here there should be a hierarchy of tasks corresponding to the hierarchy of systems, that is, the superior level system should have a structure of means (their tasks) so that the aggregate of the inferior-level systems comprising it (the aggregate of their tasks) ensures the achieving of the system's goals as a whole. Secondly, the attainability of the goal depends upon the expenditure of resources on fulfilling it. A goal can be chosen where the available resources do not make it possible to create the system ensuring its attainment. For this reason, the final determination of the goals can be given only in the process of analysis as the setting of the task will change both depending upon the resource constraints which also can be revised from the results of the analysis and upon the technical possibilities disclosed in the course of the analysis (for example, the possibility of creating a multigoal system capable of carrying out a broader range of tasks instead of a specialized system).

In the process of operations research, on the basis of an analysis of the probable conditions for carrying out the operations and the system parameters in the designated period, a logical description is given and the mathematical models are formulated for the possible variations of standard operations (methods of attaining the goal).

Depending upon the specific purpose, the operations performed by aircraft systems can be divided into information (inspection, communications and so forth) and transport operations.

Depending upon the workfront, the scope of activity and the degree of involving technology and human resources, operations can be divided into volume (supervolume) and local (sublocal).

Thus, in the process of operations research there is a choice of goals and the methods of attaining them. This makes it possible to formulate the operational links and constraints which are part of the model of the criterion to assess (select) the alternative systems.

In parallel the possible technology is studied and elaborated for achieving the selected goals in the form of variations of design decisions (in the instance of the discrete positing of a problem, a comparison of a finite number of variations) or the acceptable ranges for the change in the system's parameters (in the instance of the continuous positing of the problem, the systematized sorting out of an infinite number of variations in the designated range of parameter changes).

Project or design research, like the entire process of systems analysis, is an iterative process. In designing on the basis of the tactical and technical requirements worked out by the client from the results of operations research, the system's structure is formed, the base subsystems are selected and modified, new ones are designed and the system is synthesized and analyzed.

In the event of an unsatisfactory solution, the design process is repeated. The iterations are carried out until a satisfactory solution has been obtained.

The use of base (standardized) subsystems in the system being designed is of a contradictory nature. On the one hand there is a reduction in expenditures on the designing and series production of the system, and in addition, the development time of the system is shortened, while on the other, the use of such subsystems can lower the level of the technical advancement and operational efficiency of the system.

In the conflict period, the superiority of one of the competing systems over another to a significant degree is determined by the degree of the system's degredation and reconstruction rate. This, in turn, is largely dependent upon the mass production of reserve means used for the reconstruction as well as upon the labor intensiveness of the reconstruction processes. It is not to be excluded that these considerations may be crucial in examining the designated contradiction.

The genesis of the idea of creating new means is related, as was already pointed out, to two sources: the rise of new tasks and the achievements in technical progress. In this regard the genesis of new plans and alternative design variations for the systems must be expected in organizations entrusted with the solving of new problems (the client) and the organizations directly developing the technology (the developer). Obviously only the combined activities of these organizations ensure the elaboration of alternatives which conform to the demands of the problems being solved and of scientific and technical progress.

As was already pointed out, in the BTS, as a rule, the central element of the system is the most revolutionary link. Within the system, as a rule, there are two or three generations of central elements (for example, of aircraft) having identical purpose but different efficiency levels. The system's average efficiency level as a whole will depend upon the degree of heterogeneity of the systems of the competing sides, that is, upon the proportional amount of different-generation central elements within the systems of both sides. In this regard a study of the replacement rate in the generations of central elements within the systems of a competing side should be one of the objects of systems analysis.

The choice of an alternative for the next generation of central elements should be made proceeding from the view that the incorporation of new types of central elements in the system makes the system an optimum one. A system which has been optimized under the supposition that it will include only new elements can be nonoptimal under the conditions where the system possesses central elements of several generations.

This conclusion follows directly from the so-called Bellman optimality principle. According to this principle, an optimum sequence of decisions possesses the property where, regardless of the initial state and the decision taken at the first moment, the following decisions should be optimal relative to the state arising out of the initial decision [34].

The heterogeneity of the BTS places a number of demands on the other subsystems. The basic demand is compatibility or the reserving of possibilities for compatibility with the central elements of several generations.

Thus, in the process of project or design research, alternatives are worked out for the design variations or the acceptable ranges for the changes in the system's parameters as well as the demands incorporated in the group of design links and constraints in a system of disciplining conditions in solving the problems of an economic assessment of systems.

In the process of carrying out the goal-oriented, operational and project research extensive use is made of forecasting methods and these make it possible to reduce the degree of ambituity in the notion of the future goals, tasks and possible paths in the scientific and technical development of BTS functional elements.

In the process of criterion research, on the basis of analyzing the goal orientation of the standard operations, the possible criteria are determined for evaluating the system and the mathematical model of the criterion function (the goal) and the overall appearance of the disciplining conditions (the matrix of conditions and vector of constraints) are formulated.

The carrying out of resource research entails the necessity of setting numerical perameters for the criterion function, the matrix of conditions and constraint vectors determining the group of economic ties. In the process of this research the following are determined: the resources required to implement the alternative programs of the "resources--system parameters" link as needed for an analytical description of the criterion function and disciplining conditions as well as the constraints imposed on the amount of the resources.

Resource analysis is carried out according to the types (material, labor, financial and so forth) as well as in terms of the stages of the system's life cycle and elements. In the process of this research, the necessary and sufficient degree of aggregating the forecast estimates is determined. In accord with the scope of the initial information, a choice is made of the forecasting method and the composition of the essential factors and variables determining the effectiveness of the system and characterizing the state of the stages of their life cycle is defined. The mass of statistical information is formed, it is analyzed and processed, forecast resource models are constructed and the accuracy and reliability of the forecast calculations are assessed.

The choice of alternatives for the optimization of the system is made using the selected criteria considering the formulated constraints which are determined by the operational, design and economic links.

CHAPTER 2: FORECASTING THE DEVELOPMENT OF LARGE TECHNICAL SYSTEMS

2.1. Functions and Tasks of Forecasting

The enormous impact of scientific and technical progress on the development level of productive forces necessitates the presence of constant controlling actions on the nature of scientific and technical development and the introduction of their results into the industrial production sphere. The management of scientific and technical development is an important element in production management.

Production management has gained the widest development in a socialist society. Under the conditions of a socialist economy, national economic management is an objective necessity. Public ownership of the means of production makes it possible to have control on a scale of the entire national economy. Production management is a most important function of the socialist state.

Management includes three basic elements: planning, the organization and management itself (or control) of production. These management elements are interrelated and interdetermined and represent a single process, a single management system. The primary element is planning which determines the production development goals. The organizational structures and procedures are formulated for the established goals. Within the set structures and procedures production is controlled under the interests of attaining the goal.

In keeping with the development of the productive forces and the accelerated base of scientific and technical progress, the role of management has grown and the management system has become more complex and advanced. Production planning and primarily long-range planning and forecasting have assumed particular significance.

Planning as an element of management is an information process. A particular feature of this process is the presence of a time shift of the information output in relation to the information input. In planning the information flows on the past (retrospective information) are the input while the flows of information on the future (prospective information) are the output (Fig. 2.1).

Along with retrospective information, in adopting plan decisions, information is also used on the state of the planning object and the environment (background) at the moment the plan is worked out; this is information on the present. In relation to the planned period, this information is also information on the past and for this reason it can be conditionally classified with the retrospective information (conditionally retrospective information).

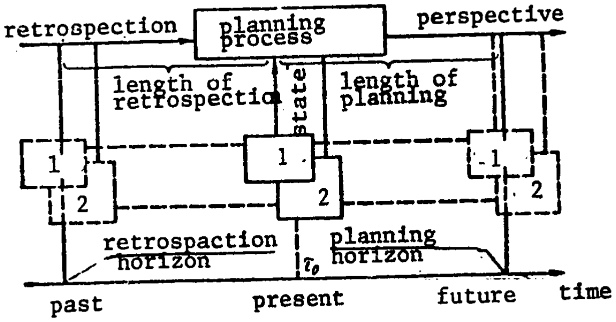


Fig. 2.1. Characteristics of information inputs and outputs of planning process

Key: 1--Object; 2--Background

The amount of the time lag of the information output and input of the process depends upon the length (the lead time) of planning, that is, upon the time interval in the future for which the plan is worked out. The greater the amount of the lead the more needed the length of retrospection[1] and, consequently, the greater the time lag between the information input and output of the planning process. Depending upon the lead time, four stages are distinguished in national economic planning.

1. Operation calendar planning (with a lead time from an hour to a month).
2. Current technical and economic planning (up to 1 year).
3. Perspective and long-range planning (up to 15 years).
4. Forecasting.

The planning stages are oriented not only in time but also in the planes of the functional and territorial articulation of the planning object. The scope of the functional and territorial levels of the hierarchy by the planning stages (the space of their functioning) varies.

Forecasting and long-range and perspective planning encompass the superior levels of the functional and territorial hierarchy: from the national economy down to the enterprise. Operational-calendar planning encompasses the inferior levels of the hierarchy on the planes of the functional and territorial articulation of the

---

[1] The length of retrospection is the time interval of the object's functioning in the past (from the retrospection horizon to the present) for which the necessary and sufficient retrospective information is available. The retrospection horizon is the name given to the most distant point in the past on a time scale at which there is the necessary and sufficient information.

planning object, from the work area to the enterprise. Current technical and economic planning holds an intermediate position between them.

The designated scheme for the scope of the hierarchy levels in terms of planning stages is continuously being transformed. Under the impact of the extensive use of computers in planning, the planning stages with a short lead time (operational-calendar and current technical-economic planning) are encompassing ever-higher levels of the hierarchy. At the same time, under the impact of the accelerated pace of scientific and technical development, the stages with a longer lead time (perspective and long-range planning and forecasting) are encompassing the ever-lower levels of the hierarchy and the planning horizon[2] is being widened.

Planning can be divided into two stages: forecasting and planning per se including the first three stages and termed the plan elaboration stage. Direct links between these stages occur at the boundary of long-term planning and forecasting. They have a common sphere of application in the functional and territorial planes and an identical scheme of information flows.

The fundamental distinction between planning per se and forecasting is the nature of the output information, that is, the directive nature of planning information (plan--directive) and the orientation or guideline nature of forecast information (forecast--orientation). These differences are caused by the significant reduction in the accuracy and reliability of the information produced on the future with an increase in the depth or length of planning.

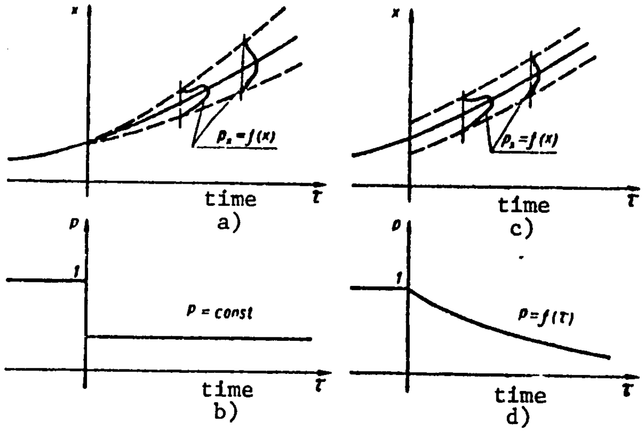


Fig. 2.2. Dynamics of the confidence interval for assessing parameters of a planning object

---

[2] The planning or forecasting horizon is the most distant point in the future on the time scale at which the state of the planning object is assessed.

Fig. 2.2 gives a diagram for the change over time in the confidence interval for assessing the state of a planning object for one parameter x. If the state of the object is described by the set of parameters $\{x_i\}$, then the curves for $x(\tau)$ given in the graphs are transformed into planes of the $(n+1)$ order, where n--the number of parameters describing the state of the object.

In constructing the graphs (Fig. 2.2, a) it is conditionally assumed that the retrospective information on the object is fully reliable. Mistakes of measurement and mistakes arising out of interference in the data transmission channels in this instance are not considered.

However, here it must be pointed out that the value of older retrospective information is reduced, its predictive force is lowered, that is, there is a discounting of retrospective information. The older the retrospective information is, the fewer the germs of the future and the greater the vestiges of the past.

The perspective information worked out in planning is of a probability nature and has a certain reliability within the limits of the confidence interval. An increase in the length of planning leads, with a constant confidence probability (Fig. 2.2, b) P = const., to a widening of the confidence interval of the estimate (in Fig. 2.2 a, it has the form bounded by the diverging curves). With a constant value of the confidence interval (the confidence interval is bounded by the equidistant curves, see Fig. 2.2, c), the confidence probability $P = f(\tau)$ is reduced (see Fig. 2.2, d). Thus, the reliability and accuracy of perspective information produced in planning is substantially reduced with an increase in the length of planning, that is, there is a discounting of the perspective information.

A directive nature cannot be ascribed to perspective unreliable information, but this information indicates a probable state of the future and is a guideline for future planning decisions with a shorter length of planning. Forecasts even with a relatively small degree of reliability make it possible to reduce the uncertainty of our knowledge about the future and, consequently, lower the risk of the present planning decisions and the harm from their nonoptimality which can arise beyond the planned period.

As we see, the time factor is primary in delimiting the concepts of forecasting and planning per se (the elaboration of a plan). It determines the limits of the processes of planning per se and forecasting. The length of forecasting theoretically is not limited. In practical terms it is obviously advisable to set such limits proceeding from the necessary and sufficient reliability of the estimates for the state of the planning object in the future. Thus, in terms of the lead time, forecasting holds the superior level in the hierarchy and then comes the elaboration of plans.

Let us give the basic concepts of forecasting theory: the forecast, forecasting and prognostics [18]. A forecast is a probability judgment concerning the state of a certain object (process or phenomenon) at a certain moment of time in the future and (or) the alternative ways of achieving them. Forecasting is the process of formulating the development forecasts of an object on the basis of analyzing its development trends. Prognostics is the science studying the patterns of the forecasting process.

In certain sources concepts are encountered which replace the concept of forecast: ing: prediction and foresight. Prediction is a reliable judgment based on a logical sequence concerning the state of a certain object (process or phenomenon) in the future. Prediction is the advance reflection of reality based on a knowledge of the development laws of an object (process or phenomenon).

Prediction and forecasting differ in terms of the reliability of the future assessments and foresight is a broader generic concept which includes both of the previous ones. Thus, the logical formulas for the different types of processes of producing information on the future (foresight) can be written thus: forecasting--"it probably will be," prediction--"it will be" and planning--"it should be."

The concept of futurology as a science dealing with the future has become widespread in foreign terminology. Being in a certain sense the equivalent of the term "prognostics," the concept "futurology" significantly and unjustifiably broadens the subject of the science, making it all-encompassing and including all aspects of the problem of the future.

There are also other viewpoints on the question of the relationship between the concepts of forecasting and planning. At times an opposition is ascribed between forecasting as the foreseeing of spontaneous uncontrollable socioeconomic processes characteristic of capitalism and planning as the defining of development trends in the future for controllable processes in society and the national economy under socialism. Such an approach to the forecasting of national economic development is invalid, as forecasting and planning (the stage of plan elaboration) have the same informational and socioeconomic nature.

In other instances the nature of the output information is considered to be the primary factor in delimiting the concepts of planning and forecasting, that is, the directive nature of the plan and the noncompulsoriness of a forecast. This distinction between plans and forecasts is secondary and is caused, as was already pointed out, by the time factor and the related greater level of forecast ambiguity. The supporters of this view feel that the plan and the forecast of national economic development can be compiled for the same period. Obviously the presence of two sets for the same future period--a directive uniform indication and a noncompulsory multivariant guideline--are merely capable of misleading production and depriving a production collective of a unity of goal.

There is the viewpoint that forecasting is the preplanning studies, that is, the process preceding planning. The unity of the tasks of planning and forecasting and the commonness of their principles and methods make it ill-advised to have a fundamental division and opposition between these concepts. There is a single production planning system as a system of producing information on the future and this includes the forecasting stage and the plan elaboration stage.

Thus, production planning is a unified system for generating information on the future and this system does not have a formal limitation in time and consists of two stages--forecasting and plan elaboration. With the specific features of these planning stages, they are united primarily by the commonness of the goals (producing information on the future) and the tasks.

For gaining information on the future it is essential to study the national economic laws and to determine the causes and driving forces of this development. This is the basic task of planning and forecasting. Social requirements, technical possibilities and economic advisability are the basic driving forces in the development of production. In accord with this it is possible to point to three basic tasks for planning and forecasting: the setting of the national economic development goals, the seeking out of the optimum ways and means for achieving them and determining the required resources for attaining the set goals.

The choice of goals is the result of analyzing the sociopolitical tasks which must be carried out in a society and which reflect the objective action of the economic laws of socialism.

The selection of goals is preceded by the elaboration of alternative goals, by the constructing of an hierarchical system or "tree of goals," by the ranking of the goals and the choice of the leading links. The initial prerequisites for goal selection are, on the one hand, the real possibility of solving the given alternative and, on the other, its optimality in terms of the efficiency criterion.

The next task of planning is to study the possible ways and means of attaining the set goals. The ways and means of attaining the goals are determining on the basis of analyzing the development of the national economy and scientific-technical progress. Here in the forecasting process there is a restricting of the area of alternative ways and means for achieving the set goals, that is, the area of optimum decisions is defined. The sole alternative criterion which is optimal in terms of the accepted vector is determined in the process of working out the plan.

It must be pointed out that, depending upon what task is carried out first, two types of forecasting are recognized: research (or exploratory) and normative. The research or exploratory forecasting is the name given to the drawing up of forecasts for objectively existing development trends on the basis of an analysis of historical trends. This type of forecasting is based on the use of the principle of development inertia where the forecast is oriented in time "from the present to the future." A research forecast is the picture of the forecast object's state at a certain moment in the future as obtained as a result of examining the development process as movement by inertia from the present to the forecast horizon.

The forecasting of the development trends of the forecast object where these trends should attain certain sociopolitical and economic goals at a set moment in the future is termed normative. In this instance the time orientation of the forecast is "from the future to the present."

The discrepancy of the normative and research estimates of a forecast object at a given moment of future time is the consequence of the "need--possibility" contradiction. A composite forecast is based on the research and normative forecasts.

The choice of the goals and means for attaining them without fail should be combined with setting the resource requirements. In setting the resources it is essential to view the planning and forecast resource matrices (financial, labor, material and energy) as well as the production capacity and time resource matrices. Also assessed are the required resources and the probable constraints on their amount within the range of the lead time of the plan or forecast. The forecast's resource matrices are the most important initial data in drawing up the national economic balances in long-range planning.

The driving forces of development do not operate in isolation, they are interrelated and interdetermined and can be shown in the form of a connected triangular graph (Fig. 2.3). The vertexes of this "causal triangle" identify the driving forces of production development and its edges are the two-way ties between them. For this reason the tasks of planning and forecasting cannot be viewed separately. In the process of forecasting and plan elaboration without fail there is an analysis of the interaction of the goals, the methods and technical means of attaining them and the required resources for realizing them and using the accepted efficiency criteria the optimum national economic development paths are determined.

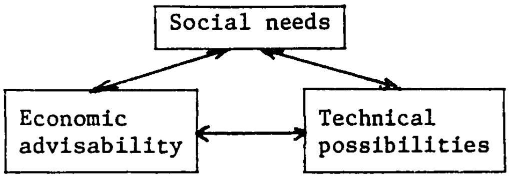


Fig. 2.3. Triangular graph of driving forces of development

Regardless of the commonness of the tasks, their positing in forecasting and planning differ. In planning there is the following scheme: goal--directive, the ways and means of achieving them are determined while resources are limited. In forecasting the scheme is different: the goals are theoretically attainable, the ways and means of attaining them are possible while the resources are probable.

As is seen, the plan will contain only one (optimum) solution while the forecast will have a range of alternatives. This particular feature is also a consequence of the time factor as the large time lead causes a high degree of ambiguity in the information on the future and, consequently, a widening of the confidence interval of the forecast estimates (the probability nature of the estimates). The tasks of planning and forecasting also differ in terms of the breadth of coverage. While the tasks of forecasting are global ones, the tasks of the other stages of planning are tasks of a lower rank. Thus, the global goal of forecasting national economic development in the USSR--the creation of the material and technical base of communism--is transformed in the Tenth Five-Year Plan as a more concrete goal of a lower rank. "The main task of the Tenth Five-Year Plan," as is pointed out in the Basic Directions for USSR National Economic Development in 1976-1980, "is to consistently carry out the communist party's course of raising the material and cultural standard of living of the people on the basis of the dynamic and proportional development of social production, a rise in its efficiency, the acceleration of scientific and technical progress, the growth of labor productivity and the greatest possible improvement in the quality of work in all the national economic units" [1]. The goals of the current plans are defined in accord with the main task of the five-year plan. The aim of each inferior planning level is to ensure the achieving of the goal by the superior level, that is, a compatibility of goals among the different planning levels should be achieved in planning.

National economic planning is carried out on a basis of the conscious use of the law of planned, proportional national economic development and is correlated with the basic economic law of socialism. Marxist-Leninist economic science is the scientific basis of planning theory.

In our nation national economic planning has been carried out from the first years of the founding of the Soviet state. In 1920, under the direct leadership of V. I. Lenin the GOELRO [State Commission for the Electrification of Russia] Plan was worked out. This first forecast plan for the socialist reorganization of the Soviet republic's national economy through large-scale machine industry and electrification was designed for 15 years.

In the following years a number of other forecasts was worked out. In 1920, under the leadership of the Soviet scientist G. S. Strumilin, a demographic forecast was worked out and this was a forecast for the size of our nation's population for 1920-1941. Prior to the Great Patriotic War, under the leadership of T. S. Khachaturov, a forecast was drawn up for the development of transportation for 10-15 years. In 1945,-1946, the USSR Gosplan drew up a national economic development forecast, in 1948, a plan for the transformation of nature, in 1959-1960 a general perspective of national economic development for a 15-year period (up to 1975) and then for 20 years, up to 1980. In 1967-1969, a plan was elaborated for the development and location of the productive forces up to 1980.

The long-range five-year plans compiled considering these forecasts played an important role in the development of the socialist national economy. At present the initial projections of national economic development are being worked out for a 15-year period (1976-1990) and the forecasts up to the year 2000.

Starting in the 1950's, in a number of the capitalist nations, and primarily the United States, a great deal of attention has been devoted to forecasting and its science and attempts have been made to compile development plans (programs).

However, the political and economic structure of a capitalist society which is determined by the private ownership of the means of production and by capitalist production relationships creates an objective impossibility of effective management, planning and forecasting of production development. Characteristic of a capitalist economy are long-term studies only for individual, relatively stable sectors (chiefly for the military sectors) while unsuccessful attempts have been made to unify their results in the so-called macroeconomic structure.

The deviating of the forecasts and plans from actual reality in the capitalist world is caused by the discrepancies between the particular patterns characteristic of the individual industrial complexes and the social conditions of their manifestation.

The financial crisis which has engulfed many capitalist nations in recent years has again convincingly demonstrated the impossibility of efficient economic management in the capitalist nations, including its highest form, state monopolistic capitalism.

2.2. A Classification of Forecasting Methods

In the modern literature on forecasting a good deal of attention is devoted to the questions of classifying the forecasting methods. At present one could count more than a score classification systems of various authors. However, up to the present there has been no unified classification of forecasting methods which is useful, sufficiently complete and open (in the sense of a possibility of broadening).

Probably, prognostics as a science has not yet reached a development level where it would be possible to create a unified classification, and for this reason it is not the aim of the given section to conpensate for this shortcoming as a whole. Here an attempt has been made to formulate the goals of a classification, to examine the possible ways of attaining them and to review certain examples from the past.

What are the aims in classifying forecasting methods? Two basic aims could be mentioned. In the first place, there is classification for the purpose of studying and analyzing the methods and, secondly, classification for the purpose of selecting a method in working out the forecasts of the object.

As is known, there are two basic types of classification: successive and parallel. A successive classification presupposes the separating out of particular groups from more general ones. This is a process which is identical to the dividing of a generic concept into specific concepts. Here the following basic rules should be observed; the basis of the division (the feature) should remain the same in the formation of any specific concept; the groups of the specific concepts should exclude one another (the demand of the absence of overlapping classes); the groups of specific concepts should exhaust the group of the generic concept (the demand of the full coverage of all objects of classification).

The parallel type of classification presupposes a complex basis of classification consisting not of one but rather of a whole series of features. The basic principle of such a classification is the independence of the selected features each of which is essential, all of them together are simultaneously inherent to the subject and only their aggregate provides an exhaustive idea of each class.

A successive classification can be given a visual interpretation in the form of a certain geneological tree and for this reason makes it possible to encompass the entire area of classification as a whole and determine the place and relationships of each class in the general system. It is more accessible for the purposes of study and makes it possible procedurally to represent the classified area of knowledge in a more orderly manner.

In the parallel system of classification, each class can be interpreted as a certain area in the n-dimensional space of classification features. This interpretation, naturally, is less visual and procedurally is not as convenient for presenting and studying the classes. However, the classification possibilities of such a system are greater than the successive approach, since the complex specific features make it possible to provide a more detailed classification with no overlapping of the classes. For this reason, in practice, for example, in the process of selecting a class of methods for an object characterized by the given set of parameters, this classification is more effective than the successive one.

Examples of both types of classifications are known among the classifications of forecasting methods. Certain authors prefer to give a classification of forecasts and others a classification of the forecast objects. Let us give certain examples of classifications and briefly describe them.

The classification of G. M. Dobrov [13] gives extrapolation, expert estimates and modeling as the basic classes for forecasting methods. Then follows a level of 8 types and below that 19 generalized names of methods.

In examining this classification one can note a violation of the principles of an ideal classification in it. On each level there is not a unified classification feature and the demand of an absence of overlapping types is not observed (the types of the modeling class can be partially put among the expert methods while types of the extrapolation class are partially among the mathematical models). On the inferior level such narrow specific methods as an interview are given simultaneously along with the almost all-encompassing groups (incidentally also overlapping) such as mathematical economics models and probability statistics models.

The classification of E. Ianch [50] on the upper level cannot be reduced to a single feature: 1) intuitive methods, 2) methods of exploratory forecasting, 3) methods of normative forecasting, 4) methods with feedback. As we can see, the second and third classes are determined by the aim of forecasting while the first and fourth are determined by its apparatus. The overlapping of the types is also apparent. Thus, the Delphi method from the first class uses the feedback principle with an expert, that is, it could be put in the fourth class, the writing of a senario (the second class) usually preceeds the constructing of the tree of goals, that is, it is included in a normative forecast (third class) and so forth.

V. A. Lisichkin [49] gives a system of features for forecast classification. This system is a parallel one for 18 features, one of which is a method used for the forecasting.

Thus, this is not a pure classification of methods but rather a mixed classification in terms of the types of objects, goals, the tasks of the forecast and the methods of carrying it out.

Here are the classification features: the nature of the forecast object; the scale of the forecast object; the number of forecasted objects; the nature of the link of the forecasted object with other objects; the nature of the change process in the object; the lead time of the forecasted event; the degree of localizing the forecast on the scale of probable situations; the method used for forecasting; the number of methods used for forecasting; the nature of the process of compiling the forecast; the relationship of the predictor to the forecast object; the system of knowledge underlying the forecast; the form of expressing the forecast's results; the goal of the forecast; the purpose of the forecast; the degree of understanding and soundness of the forecast; the method of testing the reliability of the forecast; the area of science the object of which is being forecast.

The given mixed system of classification features provides an opportunity to put each of the forecasts in a relatively narrow class and describe it from various aspects, however the utility of such a classification is not very clear from the viewpoint of the two possible aims formulated above.

In the work [23], this system of features has been developed and presented in the form of an hierarchical structure. In this form it is more complete and finished. Here four basic aspects of classification are examined: the process of making the forecasts, the object of forecasts, the predictor and the forecasting method.

In the classification by the aspect of the forecasting method, the forecasts are divided into two groups: those based on unsystematized knowledge (everyday) and those employing a system of scientific knowledge (scientific). The latter are divided into hypothetical, theoretical and empirical. Each of these classes is divided in terms of the employed type of methods into general scientific, interscientific (the method uses the apparatus of a specific science) and special scientific (methods used only in a narrow area of science). The classification is then made in terms of the number of methods employed in making the forecast. If there is one then it is a simplex forecast, if there are two methods it is a duplex, and if three or more then it is a compound forecast. Then follows the feature of the time lead by which forecasts are classified into long-term, medium-term and short-term. Finally, there is the feature of forecast accuracy by various scales: by the scale of probabilities, by the scale of parameters and by the semantic scale.

Without going into a detailed analysis of this classification, we would point out that the presenting of it in the form of a polyhierarchical tree has advantages from the viewpoint of the procedure for expounding the entire range of forecasting problems. At the same time, in essence, it remains a parallel-type classification, however it does not solve and does not ease the problems of the choice of forecasting methods.

An attempt to approach the problem of selecting the forecasting method on the basis of a classification of information data on the forecast object has been undertaken in [48]. In accord with the data defined by the classifier, the initial information is assigned a certain eight-digit code which is then compared with a table of the known forecasting methods. Appropriate methods are selected in the process of this comparison. The data classification features by categories are given below.

Quantitative Data	Qualitative Data
Random--Nonrandom	Single-factor--Multifactor
Singular--Mass	Homogeneous--Heterogeneous
Discrete--Continuous	Scalable--Unscalable
Periodic-Nonperiodic	Cyclical--Trajectory
Stationary--Nonstationary	Stationary--Nonstationary
Reliable--Unreliable	Reliable--Unreliable
Representative--Nonrepresentative	Representative--Nonrepresentative

Each place in the code can contain a 0, 1 or 2 and the 2 is used in the instance that the given place is indifferent to the forecasting object or method. The table of methods gives 38 names and their corresponding information codes.

It should be pointed out that the very idea of selecting a method according to the information possessed about the object is extremely enticing, however the realization of this idea cannot be considered satisfactory. In the first place, the given classifier has not been sufficiently worked out as the concepts of reliability, representativeness, cyclicalness, trajectoriness and stationariness for the quantitative data are unclear. Secondly, in order to identify the method it is not enough to consider just the initial information but rather it is also essential to possess information on the nature of the object, the goal and the tasks of the forecast, as well as the demands which are made upon the forecast's quality.

Even in the instance that all the listed data are present, the process of selecting the method remains a creative, unformalized process which should not and cannot be replaced by the mechanical substituting of numbers at least at the present development level of prognostics.

From the examination of the above-given classifications of forecasting methods it is apparent that at present this problem cannot be considered satisfactorily resolved.

In our view, at present it is impossible to present a unified classification which would satisfy both aims formulated at the start of the given section. For this reason we propose two classifications: the first of the successive type for the purposes of visual presentation and an analysis of the methods and a second of the parallel type for the purposes of facilitating the choice of the method for the specific forecasting object.

Fig. 2.4 shows the first of the designated classifications [35]. On the first level, the classification feature (according to the information basis) divides the methods into two classes: factographic and expert. The factographic methods are forecasting methods which use as the information base real facts that occurred in the past. These facts can be recorded on any information carrier and have both a quantitative and qualitative nature. In opposition to the factographic methods, the expert methods are based upon the processing of opinions and judgments by specialists or experts in one or another area of knowledge and these are obtained in the process of various specialized procedures for their collection.

The classification feature of the second level has been formulated as the method of employing the information about the object. In the class of factographic methods, three types have been established for this feature.

The first type is the aggregate of extrapolation and interpolation methods. Characteristic for this type of methods is the use of initial information for constructing fitting functions. Then the value of the segment of lead time is substituted in the found dependence and the obtained value of the function for the sought forecast is set.

The second type of factographic methods is based upon a study of the relationships between two or more variables in the forecast object with the subsequent determining of the future values of some variables using the values of others which are known or have been determined by other methods. The apparatus of multidimensional mathematical statistics underlies such methods. In this type of methods a specific place is held by the so-called lead methods which are based on a study of relationships between the scientific-technical information and the scientific-technical progress.

As a rule, for all the methods of this type, after elucidating the relationships between the variables and the constructing of one or another statistical model, the

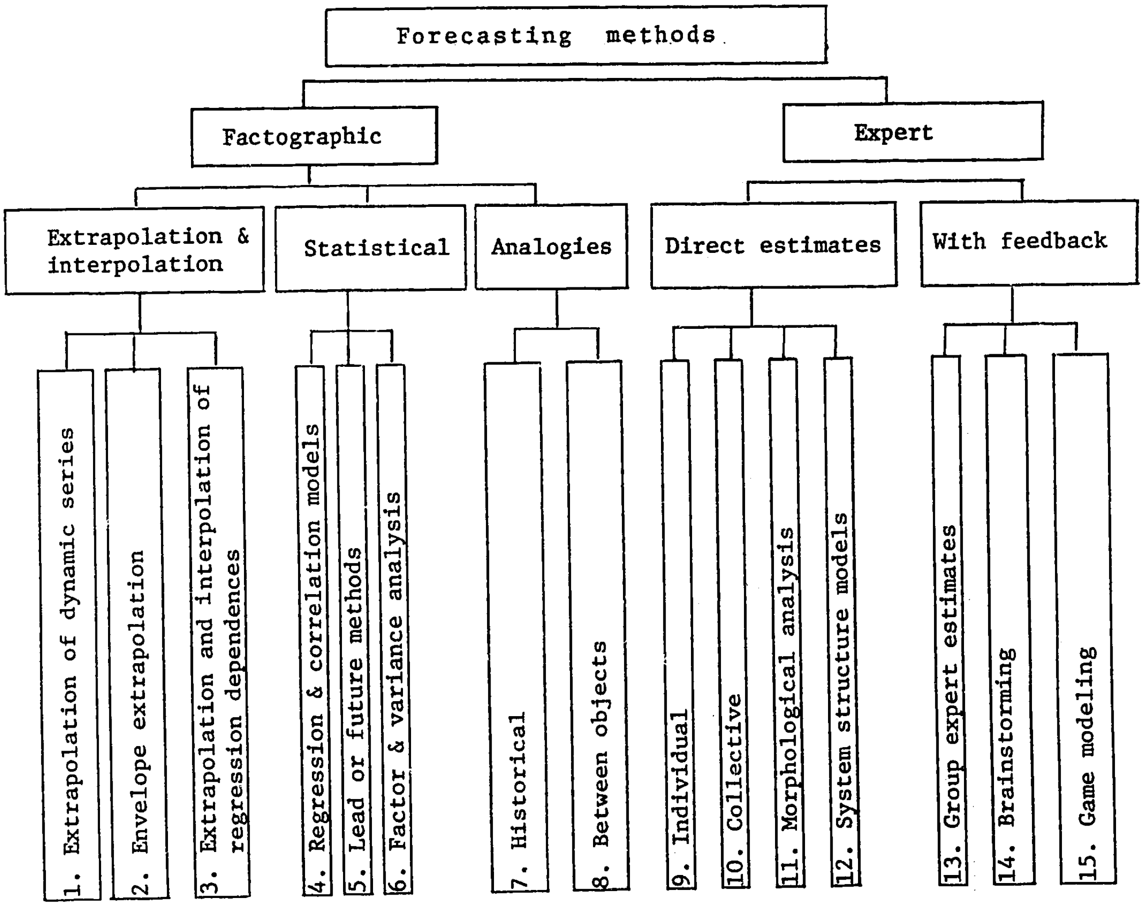


Fig. 2.4. Classification of forecasting methods

actual forecast is obtained by the extrapolation or interpolation of the dependent variables in relation to the independent variables.

The following type of methods is based upon a study of the future development of certain objects following the development patterns of their analogous objects. Here it is possible to use both quantitative and qualitative information. Moreover, within this type one can isolate the research on analogies between objects of the same nature but having a certain historical break in the development level and analogies between objects which differ in their nature. In the first instance, for example, this could be two countries standing at different levels of social and technical development, and in the second, analogies known from foreign sources of literature between the development of a biological system and the spread of a new production method and so forth. It is essential to point out that in the practice of scientific-technical and economic forecasting, particularly in our nation, this type of method is applied in practice comparatively rarely.

In terms of the method of using the information obtained from expert specialists, there are two types of methods: direct estimates and those with feedback. Here the difference is that in the first instance the expert information obtained is processed and given directly as a result. In the second type of methods, the result is obtained in a process, as a rule, of several approximations and at each step the results of processing the previous one influenced the experts, that is, there is feedback with the experts.

The inferior level of our classification is comprised of comparatively narrow groups of methods which are close in their essence.

It is essential to point out that the presented classification encompasses only simple (singular) methods[3] and does not include the composite ones. As a rule, composite methods of varying complexity are employed in practice. For example, the PATTERN method [25] includes as component elements the following: the writing of a scenario, morphological analysis, the constructing of a tree of goals and collective expert estimates. The matrix forecasting method includes the construction of a model graph of the object and collective expert estimates. The patent methods ordinarily employ the construction of a certain classification tree for assessing the importance of patents, the methods of statistical analysis for examining the statistics of patenting and its internal and external relationships, the methods of extrapolation and analogies of patenting dynamics and introduction dynamics.

Obviously with the development of the apparatus of prognostics, the number of singular methods will grow and at the same time the complexity of the comprehensive methods will rise with an increase in the demands made upon the quality of the forecasts. In this regard, as was pointed out above, at the present level in the development of prognostics it is hard to propose a classification of the parallel-type forecasting methods which would make it possible to select uniformly a method for forecasting a certain range of parameters in an object. It is only possible to list groups of preferential methods which are employed under various conditions and for different objects.

In selecting a forecasting method, it is essential to consider the following basic factors: the type of forecast to be worked out; the volume and type of initial information about the object; the ratio of the base time τ_0 and the lead time τ_y of the forecast.

In terms of the first factor, it is possible to isolate exploratory and normative forecasts. Let us designate the first by the digit 0 and the latter by the digit 1.

For the second factor, the information one, it is possible to isolate four forecast subgroups: 0--there is a very limited amount of information about the object; 1--there is qualitative information; 2--there is statistical information about the object or estimates of the basic probability characteristic; 3--there is determined quantitative information.

---

[3] By singular methods one understands methods which are not broken down in a certain sense into other, simpler ones based, as a rule, on a certain type of information and employing a specific area of mathematical procedures.

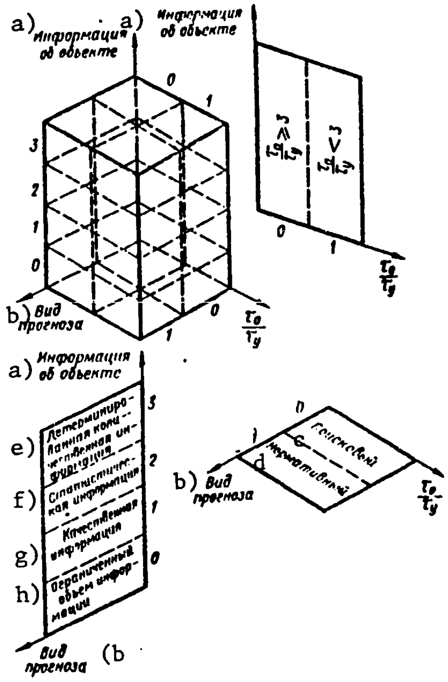


Fig. 2.5. Space of forecast classification features

Key: a--Information about object; b--Type of forecast; c--Exploratory; d--Normative; e--Determined quantitative information; f--Statistical information; g--Qualitative information; h--Limited amount of information

For the third factor which cannot be examined for all the number combinations of the first two, it is possibl- to isolate the following forecast subgroups: 0--with a ratio of τ_0/τ_y 3; 1--with a ratio of τ_0/τ_y 3. The figure 3 has been taken as a result of generalizing the opinions of a number of authors on the minimum acceptable ratio of the base time to the lead time for a forecast. In their majority these relate to the extrapolation methods of exploratory forecasting.

For normative forecasting, it would be correct to formulate the gradations for the third factor depending simply upon the lead time: short-term, medium-term and long-term forecasts. For maintaining the commonness of the classification, we will keep the two gradations for the normative forecasts as well using the digit 0 for the short- and medium-term forecasts and the digit 1 for long-term ones.

Thus, we have a three-dimensional space of features in which there are 16 areas corresponding to the various values of their codes (Fig. 2.5). For each of these areas it would be possible to name several methods from those listed above (see Fig. 2.4) which would preferentially be used under conditions corresponding to the value of its coordinates. The simple methods ordinarily are incorporated in the comprehensive forecasting methods for this area. It is possible to give the following tentative division of methods, using the numeration of the methods shown in Fig. 2.4 for the following areas:

Area Code	Number of Preferential Methods	Area Code	Number of Preferential Methods
000	8, 9, 10, 13	100	9, 10, 14
001	9, 10, 13	101	9, 10, 13, 14
010	7, 8, 9, 10, 13	110	9, 10, 14, 15
011	9, 10, 13	111	9, 10, 12, 13, 14
020	3, 4, 5, 6	120	4, 5, 6
021	4, 5, 6, 9, 10, 13	121	4, 5, 6, 10, 13
030	1, 2	130	10, 11, 12, 14, 15
031	5, 7, 8, 9, 10, 13	131	11, 12, 13, 14, 15

The given parallel classification can be used as a guideline in the problems of selecting suitable methods. The solution to this problem as a whole remains a complex unformalized process as was already pointed out in the given paragraph.

2.3. Expert Forecasting Methods

The area of employing expert forecasting methods is the scientific-technical objects and problems an analysis of which either completely or partially cannot be submitted to a mathematical formalization. These methods provide an opportunity to construct an adequate model of future scientific-technical development on the basis of opinions of persons working in science and technology (experts).

The use of expert opinions as sources of information about the forecasted object is based upon the hypothesis that they possess ideas about the ways to solve particular or global problems, apriori judgments about the importance of different decisions and intuitive guesses about alternative and possible variations for the development of the studied object.

Let us examine the most widely found methods of expert evaluation which are employed in scientific and technical forecasting.

2.3.1. Individual Expert Estimates

Individual expert methods are based on the use of the opinions of experts who are specialists in the appropriate specialty, independently of one another. The most frequently employed are two methods of making a forecast: on the basis of a conversation between the forecaster and the expert following a previously elaborated program (the interview methods); on the basis of the extended and careful independent work by the expert on posed questions (the analytical estimate method).

The interview method is the simplest method of expert evaluation and here the specialist's opinion is elucidated by an expert. The analytical estimate method, on the contrary, provides an opportunity for the expert to use all the information needed by him about the forecast object. The expert draws up his ideas in the form of a report.

The basic advantage of the designated forecast methods is the possibility of making maximum use of the individual abilities of the expert. However, these methods are little suitable for forecasting the most general strategies due to the limited knowledge of one expert specialist about the development of related areas of science.

2.3.2. Collective Expert Estimates

The method of collective expert estimates is based on the principles of elucidating the collective opinion of experts on the development prospects of the forecasting object.

The use of these methods is based on the hypothesis of the expert's ability with a sufficient degree of reliability to assess the importance and significance of a scientific and technical problem, the prospectiveness of developing a certain area of research, the time for completing one or another event, the advisability of selecting one of the alternative development paths for the forecast object and so forth.

The advantage of these methods is the possibility of exchanging opinions between the expert specialists, the orientation of the ideas toward the strategic goals and the use of internal and external feedback in the heuristic process. The basic shortcomings of these methods are the possibility of an influence of the authorities and the opinion of the majority, the difficulty of a public abandoning of one's viewpoint and so forth.

At present, expert methods based on the use of special commissions have become widespread and here the groups of experts at a "roundtable" discuss one or another problem in the aim of coordinating their opinions and working out a uniform opinion. This method has a drawback in the fact that the expert group in its judgments is basically guided by the logic of compromise.

In contrast to the commission method, in the Delphi method, instead of a collective discussion of one or another problem, there is an individual questioning of the experts ordinarily in the form of a questionnaire for elucidating the relative importance and dates for the occurrence of hypothetical events [50]. Then the questionnaires are statistically processed, the collective opinion of the group is formed, the arguments in favor of various judgments are generalized and all the information is provided to the experts. The participants of the expert evaluation are requested to review the estimates and explain the reasons for their disagreement with the collective judgment. This procedure is repeated several times (three or four times). As a result, the range of estimates is narrowed. The drawback of this method is the impossibility of eliminating the influence which the organizers of the questionnaires have on the experts in drawing up the questionnaires.

As a rule, the basic questions in drawing up the forecast using an expert collective include: the formation of a representative expert group; the preparations for and carrying out of the expert evaluation; statistical processing of the obtained results. The basic rules for solving these questions will be examined below.

Forming the representative group of experts. In forming the expert group, the basic questions are to determine the quantitative and qualitative membership of the group.

The selection of experts starts by determining the areas of scientific, technical, economic and administrative questions which are involved in solving the given problem; then lists of persons competent in these areas are drawn up.

For obtaining a qualitative forecast, a number of demands are made on the participants in the expert evaluation, the main ones being: a high level of general erudition; profound special knowledge in the area to be evaluated; the ability to adequately depict the development trends of the studied object; the presence of a psychological set for the future; the presence of an academic scientific interest in the question being studied with the lack of any practical self-interest as a specialist in this area; the presence of production and (or) research experience in the designated area.

A questionnaire is used to determine to what degree the potential expert meets the listed requirements. The method of the expert's self-evaluation of his competence is frequently employed in addition to this. In a self-assessment the expert determines the degree of his knowledge on the question being studied also using a questionnaire. The processing of the questionnaire data provides an opportunity to obtain a quantitative assessment of the potential expert's competence using the following formula:

$$K = 0{,}5\left(\frac{\sum_{j=1}^{m}\gamma_j}{\sum_{j=1}^{m}\gamma_{j\,\max}} + \frac{\lambda}{p}\right), \tag{2.1}$$

where γ_j--the weight of the gradation given by the expert for j ($j = 1, \ldots, m$) characteristics in the questionnaire, number of points;
$\gamma_{j\,\max}$--the maximum weight (the scale limit) for j characteristics, number of points;
m--the total number of competence characteristics in the questionnaire;
λ--the weight of the group marked by the expert in the self-evaluation scale, number of points;
p--the limit of the expert self-evaluation scale, number of points.

It is rather difficult to set an optimum size for an expert group. However, at present a number of formalized approaches to this question has been worked out. One of them is based upon setting the maximum and minimum size limits of the group. Here they proceed from two conditions: 1) the high average competence of the expert groups; 2) the stabilization of the average assessment of the forecasted characteristics.

The first condition is used to determine the maximum size of the expert group n_{max}:

$$cK_{\max} \leq \frac{\sum_{i=1}^{n} K_i}{n_{\max}}, \tag{2.2}$$

where c--constant;
K_{max}--the maximum possible competence for the employed competence scale;
K_i--the competence of expert i.

This conditions assumes that if there is a group of experts whose competence is maximal, then the average value of their estimates can be considered "true." Voting is used to determine the constant, that is, the group is considered elected if two-thirds of those present have voted for it. Proceeding from this it is accepted that $c = 2/3$. Thus, the maximum size of the expert group is set on the basis of the inequality

$$\frac{2}{3} n_{max} \leqslant \frac{\sum\limits_{i=1}^{n} K_i}{K_{max}}. \tag{2.3}$$

Then the minimum size of the expert group n_{min} is determined. This is done by using the condition of stabilization for the average estimate of the forecasted characteristic. This condition is formulated in the following manner: the inclusion or exclusion of the expert in the group has an insignificant influence on the average estimate of the forecasted amount

$$\frac{B-B'}{B_{max}} < \varepsilon, \tag{2.4}$$

where B--the average estimate of the forecasted amount in points as given by the expert group;

B'--the average estimate given by the expert group from which one expert has been excluded (or included therein);

B_{max}--the maximum possible estimate of the forecasted value in the adopted estimate scale;

ε--the set value for the change in the average estimate in including or excluding the expert.

The amount of the average estimate is most sensitive to the estimate of an expert who possesses the greatest competence and who has set the greatest number of points with $B \leqslant B_{max}/2$ and minimal for $B \geqslant B_{max}/2$ and for this reason for testing the realization of the condition in (2.4) it is proposed that the given expert be excluded from the group.

In the literature the rule is given of calculating the minimum number of experts in a group depending upon the set (acceptable) amount of the change in the average estimate of :

$$n_{min} = 0.5 \left(\frac{3}{\varepsilon} + 5\right). \tag{2.5}$$

Thus, the rules of (2.3), (2.4) and (2.5) provide an opportunity to obtain estimate values for the maximum and minimum number of experts in a group.

The final size of the expert group is formed on the basis of the sequential exclusion of the little-competent experts and here one uses the condition $(K_{max} - K_i) \leqslant \eta$, where η--the set amount of the limit for the acceptable deviation of competence for expert i from the maximal. Simultaneously new experts can be included in the group.

The group size is set within the limits $n_{max} \leqslant n \leqslant n_{min}$.

In addition to the above-described procedures, in the methods of collective expert estimates they also employ a detailed statistical analysis of the expert conclusions and as a result of this qualitative characteristics of the expert group are determined. In accord with these characteristics in the process of carrying out the expert evaluation the quantitative and qualitative composition of the expert group can be adjusted. The methods of determining these characteristics are examined below.

Preparation and carrying out of expert evaluation. The preparations for questioning the experts includes the elaboration of questionnaires which would contain the range of questions on the forecast object. The structural and organizational set of questions in the questionnaire should be logically linked to the central task of the expert evaluation.

Although the form and content of the questions are set by the specific nature of the forecasting object, it is possible to set general demands for them: the questions should be formulated in generally accepted terms, their formulating should exclude any semantic ambiguities and all the questions should logically correspond to the structure of the forecast object and ensure a uniform interpretation.

By form the questions can be open and closed, direct and indirect. A question is called open if the answer to it is not regulated. Questions are considered closed if their formulation contains alternative answers and the expert should choose one (or several) of them. Indirect questions are used in those instances when the aim of the expert evaluation must be concealed. Such questions are resorted to when there can be no confidence that the expert, in giving information, will be totally sincere or free of outside influences which would distort the objectiveness of the answer. Let us examine the basic groups of questions used in carrying out a collective expert estimate.

1. Questions presupposing answers in the form of a quantitative estimate, that is: on the time of the occurrence of events, on the probability of the occurrence of events or for an estimate of the relative influence of factors. It is advisable to use an uneven scale in determining the scale of values for quantitative characteristics. The choice of the specific uneven scale depends upon the nature of the dependence of the forecast error upon the lead time.

2. Questions requiring an informative reply in a concise form: disjunctive, conjunctive or implicative.

3. Questions requiring an informative answer in a complete form: with an answer in the form of a list of information about the object; with the answer in the form of a list of arguments affirming or rejecting the thesis contained in the question.

These questions are formed in two stages. In the first stage the experts are asked to formulate the most promising and least elaborated problems. In the second stage from the designated problems they choose those that are fundamentally solvable and have direct bearing on the forecast object.

The procedure of carrying out an expert estimate can vary, however here also it is possible to establish three basic stages: 1) in the first stage the experts are used to clarify the formalized model of the forecast object, to formulate the questions for the questionnaires and adjust the membership of the group; 2) in the second stage the experts work directly on the questions in the questionnaires; 3) after the preliminary processing of the forecast results the experts are used for consultation on the lacking information needed for the final formulating of the forecast.

Statistical processing of expert estimate results. In processing quantitative data contained in the questionnaires, statistical estimates are determined for the forecasted characteristics and their confidence limits as well as statistical estimates for the agreement of the expert opinions.

The average value of the forecasted amount is determined using the formula

$$B = \sum_{i=1}^{n} B_i/n,$$

where B_i--the value of the forecasted amount given by expert i;
n--the number of experts in the group.

Moreover, the variance is determined $D = \left[\sum_{i=1}^{n} (B_i - B)^2\right] \Big/ n - 1$ and the approximate value of the confidence interval $J = t\sqrt{\frac{D}{n-1}}$,

where t--the parameter determined from the Student tables for the set level of confidence probability [the number of degrees of freedom $k = (n-2)$]. .

The confidence limits for the values of the forecasted amount are figured according to the formulas: for the upper limit $A_B = B + J$; for the lower limit $A_H = B - J$.

The coefficient of variation for the estimates given by the experts is determined from the following dependence $V = \sigma/B$, where σ--the standard deviation.

In processing the results of the expert estimates for the relative importance of the scientific areas, the mean value, the variance and the coefficient of variation are figured for each assessed area. Moreover, a concordance coefficient is figured and this shows the degree of agreement among the expert opinions on the importance of each of the assessed areas and paired rank correlation coefficients which determine the degree to which the experts agree with one another.

For this the importance estimates given by the experts are ranked. Each estimate given by expert i is expressed by the number of the natural rank in such a manner that the number 1 is given to the maximum estimate and the number n to the minimum. If all the n of the estimates differ then the corresponding numbers of the natural series are the estimate ranks of expert i. If there are the same estimates among those given by expert i, these estimates are given the same rank equal to the arithmetic average of the corresponding numbers of the natural series.

The total of the ranks S_j assigned by the experts to area j (j = 1, ..., m; m--the number of studied areas) is determined by the formula

$$S_j = \sum_{i=1}^{n} R_{ij},$$

where R_{ij}--the rank of the estimate given by expert i to area j.

The mean value of the total of estimate ranks for all the areas equals $\overline{S} = \sum_{j=1}^{m} S_j/m$.

The deviation of the total ranks obtained by area j from the mean value of the total ranks is defined as $d_j = S_j - \overline{S}$. Then the concordance coefficient calculated for the aggregate of all the areas approximated for the estimate is

$$W = \frac{12 \sum_{j=1}^{m} d_j^2}{n^2 (m^3 - m) - n \sum_{i=1}^{n} T_i}.$$

The amount $T_i = \sum_{e=1}^{\alpha} t_e^3 - t_e$ is calculated with the presence of equal ranks (α--the number of equal rank groups, t_e--the number of equal ranks in the group).

The concordance coefficient assumes values within the limits from 0 to 1: W = 1 means the full agreement of the expert opinions and W = 0--full disagreement.

The concordance coefficient indicates the degree of agreement in the entire expert group. A low value for this coefficient can be obtained if a commonness of opinions is lacking among all experts as well as due to the presence of opposing opinions among expert subgroups although agreement can be high within a subgroup.

For ascertaining the degree to which the expert opinions agree with one another, a paired rank correlation coefficient is used

$$\rho_{i,i+1} = \frac{\sum_{j=1}^{m} \psi_j^2}{\frac{1}{6}(m^3 - m) - \frac{1}{12}(T_i - T_{i+1})},$$

where ψ_j--the difference (for the modulus) in the amounts of the expert ranks for area j set by experts i and (i+1); $\psi_j = |R_{ij} - R_{i+1;\, j}|$.

The paired rank correlation coefficient can assume the values $-1 \leqslant \rho \leqslant 1$. The value $\rho = 1$ corresponds to a full agreement in the opinions of two experts. The value $\rho = -1$ shows that the opinion of one expert is the opposite of the other's opinion.

For determining the significance level of the values of the coefficients W and $\rho_{i,\ i+1}$ it is possible to use the χ-square criterion. For this one calculates the value $\chi^2 = \frac{12\sum_{j=1}^{m} d_j^2}{mn(m+1) - \frac{1}{m-1}\sum_{i=1}^{n} T_i}$ (the number of degrees of freedom k = m-1) and from the appropriate tables the significance level of the obtained values is determined.

2.3.3. Brainstorm Methods

Among the intuitive forecasting methods, a significant place is held by brainstorm methods. The given methods are based upon involving all the experts in an active creative process. For using these methods in forecast studies there is an opportunity to obtain productive results over a short interval of time in a situation of the creative generating of ideas with the direct contact of experts. The following brainstorm methods can be named:

a) The direct brainstorm method the aim of which is to generate as many possible new ideas for solving a problem situation;

b) The method of a destructured relative estimate;

c) The confidence group method the aim of which is to establish the agreement among a small group of participants in the method;

d) The method of inducing mental and intellectual activity the aim of which is to find the rational choice of one or another solution to a problem situation without the establishing of quantitative estimates;

e) The method of the controlled generation of ideas the aim of which is to disclose promising and original ideas to resolve a problem situation;

f) The method of stimulated observation the aim of which is to find logical solutions to the discussed problem situation with the formulated constraints;

g) The operational creativity method the aim of which is to find the sole solution for the discussed problem situation.

The basic rules of brainstorming consist in the following: 1) The statements by the participants should be terse and clear and detailed reasoning is not required; full statements can reduce the pace and impede the development of the essential and fruitful state of emotional stress; 2) skeptical comments and criticism of previous statements are categorically prohibited; 3) each of the participants has the right to speak as many times as he wishes but not sequentially; 4) the floor is given first to those who wish to comment on the previous statement; 5) it is not permitted to read without interruption a list of proposals which could be prepared by a participant ahead of time.

Additional proposals to resolve a problem situation can occur among the participants later and for this reason all proposals which occur after the brainstorming can be written down and turned over to the organizers of the session. For brainstorming it is most productive to have a group consisting of 10-15 specialists in a session of from 20 minutes to 1 hour.

The explanation for this established fact is probably that in the process of brainstorming a majority of ideas are voiced by association with the previous idea. Along with encouraging such associative thinking it is essential to provide rapid questioning of all those who desire to be heard on the previous idea. For the occurrence of associative thinking it is essential to have a certain minimum "threshold" group size which generates an aggregate of different quality viewpoints on the discussed problem situation. The answer to the question of the make-up of the group presupposes a proper selection of participants:

1) From persons of approximately the same position (degree or title) if the participants know each other (the presence of superiors intimidates subordinates);

2) From persons of different positions (degrees or titles) if the participants are not acquainted; in this instance it is essential to make each of the participants equal by giving him a number in calling on the participant in turn by number, since the group could include candidates of sciences, for example, along with academicians.

The essential condition of the participants specialization in the area of a problem situation is not required for all group members. Moreover, it is very desirable that the group include specialists from other areas of knowledge who possess a high level of general erudition and understand the sense of the problem situation. A constructive approach to fulfilling the formulated condition consists in the coordinating of the goals of the brainstorming, the forms of informing the participant of the initial information and the competence or informativeness of the participants in the area discussed (by informativeness one understands the level of special knowledge for the participant in the discussed area).

A description of the problem situation includes: the reasons for the occurrence of the problem situation; an analysis of the causes and possible consequences of the arising problem situation (it is advisable to overstate the consequences in order to more acutely feel the need for resolving the contradictions); an analysis of world experience in solving a similar problem situation (if this exists); a classification (systematization) of the existing ways of resolving the problem situation; the formulating of the problem situation in the form of a central question with an hierarchy of subquestions (a question should be sufficiently simple in its internal structure since the narrowing of the problem encourages the efficiency of brainstorming).

The leadership of the brainstorming should be assigned to forecasters who are experienced in leading scientific discussions and problem posing and who know the procedural questions and methods. If the discussed problem situation is complicated and has a narrow specialized nature, then a specialist on the discussed question should be called in as a co-chairman for directing the brainstorming. Moreover, the group should include: methodologists who are specialists in the area of prognostics and who have experience in holding sessions and processing the results; initiators who are specialists in the area of the studied problem; analysts who are highly skilled specialists in the area of the studied problem and who are capable of summarizing the past, assessing the present state of the object and the research trends on the problem; amplifiers who are specialists in the designated problem with developed deductive thinking.

All the participants in the brainstorming should possess developed associative thinking. The brainstorming process should be carried out under conditions which help as much as possible in establishing a creative mood with a maximum concentrating of attention among the participants on the discussed problem. Since an idea advanced at a given moment could previously "mentally belong" to another participant waiting for the floor, the result of the brainstorming is considered to be the fruit of the collective labor of the entire group. The most valuable are the ideas directly tied to previously voiced ideas or arising as a result of the uniting of two or several ideas ideas into one.

It is desirable that a problem situation be presented in writing beforehand and here specific questions should be raised, such as: the goal of the brainstorming; useful ideas for solving the problem; a list of factors relating to the discussed problem situation which anticipate new approaches to solving it; a list of different viewpoints on the discussed problem; a list of questions which must be answered in order to resolve the problem more rapidly and effectively; a plan for resolving the discussed situation.

The leader of a brainstorming session should organize his opening speech in such a manner as to arouse the mental susceptibility of the participants and force the participants to feel the need of doing what the leader is asking. The process of advancing new ideas occurs in the following manner. An idea voiced by one participant in the discussion gives rise to a reaction which, because of the ban on criticism, is formulated as an accompanying idea.

As a rule, at the outset of the session it is essential to have required questioning as often the leader must stir up the participants for 5-10 minutes and create an atmosphere of a free exchange of opinions. In conducting the brainstorming the leader follows the above-listed rules and in addition he should:

1) Focus the participants' attention on the problem situation, setting its limits by the specific requirements of the problem situation and the terminological strictness of the ideas voiced;

2) He should not declare any idea faults, he should not discuss and not interrupt the examination of any idea; he should examine any idea regardless of its seeming inapplicability or unfeasibility;

3) Welcome the improving or combining of ideas. He should give the floor first to those who wish to make a statement relating to the previous statement;

4) He should support and encourage participants as this is so essential to eliminate their reticence;

5) He should create an unrestrained atmosphere thereby helping to increase the activity of the participants.

The active work of the leader is presupposed only at the outset of the brainstorming. After the participants have been sufficiently aroused the process of the promoting of new ideas occurs spontaneously. In this process the leader plays a passive role in controlling the participants according to the rules for conducting a brainstorming session. It must be remembered that the more diverse and the greater the number of statements, the broader and deeper the coverage of the examined question and the greater the probability that new ideas will occur.

The stated ideas should be taped in order not to miss a single idea and to be able to systematize them for the following stage.

After the holding of the brainstorming the ideas raised in the generating stage are systematized. The systematization is carried out by the problem situation analysis group in the following stages: 1) A nomenclature list is compiled of all the stated ideas; 2) each of the ideas is formulated in generally accepted terms; 3) duplicating and complementary ideas are determined and these are then formulated in the form of comprehensive ideas; 4) the features are established by which the ideas can be unified; 5) a list of ideas is drawn up by groups; in each group the ideas are given in the order or their commonness: from the more general to the particular which complement or develop the more general ideas.

In forecasting use is also made of the method of a destructured relative estimate (the DOO method) which is an integration of two processes of an ordinary brainstorm and the destructuring of advanced systematized ideas. The destructuring of ideas is a specialized procedure for evaluating ideas for practical feasibility in the process of brainstorming, when each of the ideas is subjected to thorough criticism by the participants in the session. The basic rule in the destructuring stage is to examine each of the systematized ideas solely from the viewpoint of obstacles on the path to achieving it, that is, the participants in the brainstorming give arguments which reject the discussed idea. In the destructuring process a counter idea can be proposed which would contain an assertion about the impossibility of realizing the idea, it would formulate the existing constraints and advance a proposal on the possibility of eliminating these constraints. The structure of the counter idea, as a rule, is as follows: "This cannot be because.... In carrying out this it is essential to use...." Thus, the result of carrying out the second stage of the DOO method is the drawing up of a list of critical comments on a group of ideas or each of the ideas as well as a list of counterideas.

The leadership of a brainstorming session in the destructuring of ideas is provided by a leader who: discloses the content of the problem situation and briefly describes the groups of systematized ideas; recalls the rules for brainstorming and concentrates the attention of the participants on the need for a thorough criticism of the ideas proposed for discussion and the advancing of counterideas; he formulates the most general idea of the first group and invites the members to have their say.

The brainstorming process in the destructuring of ideas is governed by the same rules as in the idea generation stage. However, the leader gives basic attention to preventing the voicing of arguments which substantiate the discussed idea as well as encouraging proposed counterideas.

After all those who so desired have voiced their critical comments on the discussed idea, the leader proposes the next idea for examination. The leader determines whether the next idea will be a particular idea of the discussed group of ideas or will belong to a new group of ideas. In taking this decision he is guided by the considerations of maximum criticism for all ideas belonging to the discussed group. The leader can propose that the participants criticize a group of ideas all at once but this is allowable only in the instance when the number of ideas in the group does not exceed five-seven and they are simple in terms of their inner structure.

The destructuring process is repeated until each of the systematized ideas from the list has been criticized. The advanced criticisms and counterideas are tape recorded.

The third stage of the DOO method is to assess the critical comments obtained in the destructuring process in order to draw up a final list of practically applicable ideas aimed at resolving the problem situation. The evaluation of the criticisms is as important as the destructuring of the ideas since in the destructuring stage all possible constraints impeding the practical implementation of the idea are formulated while in the generation stage the knowledge of concrete conditions under which the ideas should be realized is voiced.

The processing and analysis of the destructuring results are carried out by the problem situation analysis group. The group can be supplemented by those specialists who are empowered to take decisions on carrying out the ideas (this is particularly important in those instances when decisions must be taken quickly on a multiplicity of problems all at once).

The DOO method makes it possible to find a group solution for an arising problem situation, excluding the path of compromises. A solution in the form of a single opinion is the result of the dispassionate and successive analysis of the problem. However the method does not offer the ranking of ideas in significance or the finding of an optimum way to achieve the set goal and for this reason should be complemented by a collective expert evaluation with the subsequent statistical processing of the evaluation results.

2.4. Forecasting on the Basis of the Extrapolation and Interpolation of Trends

In examining the singular forecasting methods we have not set the task of fully representing the entire list of them either in terms of composition or in terms of the depth of examining individual methods. The following considerations have underlain the choice of the composition of examined methods and the degree of detailed examination of each of them. In the first place, the practical feasibility and utility of the methods in the tasks of forecasting BTS development, secondly, how well the given method has already been described in the literature, and thirdly, the newness and promise of the method in developing the theory and practice of prognostics.

Thus, let us turn to the problem of using the mathematical methods of extrapolation and interpolation in forecast research as this is the most fully elaborated and widely used type of factographic forecasting methods.

2.4.1. The Formal and Forecast Extrapolation

For persons who are little acquainted with forecast problems, as a rule, the question arises of why extrapolation is examined in prognostics at all as it has been presented exhaustively and in the greatest detail in mathematical literature. It is merely a question of taking the formulas, substituting the initial information and obtaining the results. What results can be obtained with such an approach is clearly demonstrated by R. Ayres [46]. For example, in formally extrapolating the growth trends for the speeds of the various types of transport over the last two centuries, by the year 1990 we will obtain speeds which significantly surpass the speed of light. In an analogous manner the curve for human life expectancy after the year 2000 rushes toward infinity. The extrapolated growth trend for the explosive power of means of destruction created by man rises virtually without restriction even after 1981. One could give a number of other results of formal extrapolation which run counter to common sense.

R. Ayres defines such forecasts as excessively exalted. On the other hand, he gives examples of blinded forecasts which do not permit the researcher to predict the possible consequences of future events, the prospects of a trend, the influence of the environment and so forth. "A simple extrapolation of trends does not presuppose the understanding of factors underlying any phenomenon and it is usually enough that these (concealed) factors remain unchanged over time" [46].

The problem of formal and forecast extrapolation has been examined more closely by G. Haustein [43]. He says that it is possible to have extrapolation on the basis of the patterns inherent to a system proceeding from the existing development trends. Mathematically the optimum fitting of results to the initial data using a polynomial to a certain degree corresponds to this.

With a different variety of extrapolation, the amounts characterizing actual data are correlated to hypotheses about the dynamics of the process over the long run. The elaboration of the hypotheses is not carried out just on the basis of past development. The closeness of the theoretical and actual data is not turned into the sole criterion for the choice of the function.

Below we will distinguish between a mathematical or formal extrapolation and forecast extrapolation. By the former we understand those methods of extrapolation (and interpolation) of dynamic series whereby no use at all is made of information about the physical or logical essence of the examined process, nonformal procedures for selecting the functions are not employed while the results are not varified by any hypotheses concerning future development of the object.

The most suitable example of mathematical extrapolation is the calculating of coefficients for the polynomial breakdown of a function for a set range of values. According to the Weierstrass theorem, a function which is stable within the interval of values (a, b) can be represented in it with any degree of accuracy in the form of a polynomial. For any value of x on (a, b), there is the valid inequality $|y-f(x)| \leqslant \varepsilon$, where ε--any small positive number.

In increasing the number of terms of the polynomial in the breakdown, it is possible to increase accuracy theoretically up to the absolute coinciding of the curve with all points. In practice this seeming advantage ends up with a loss of accuracy.

As a rule, the economic processes and the processes of scientific and technical development arise out of a certain constant, steady tendency which in prognostics is usually termed a trend, and a certain random component expressed in the fluctuation of indicators around the trend. These random components can appear as the results of random fluctuations in the external or internal variables of the object as well as random measurement results. In borrowing terms from information theory, the above-described increased accuracy of approximation leads to the emphasizing of noise (the random components) and does not filter the signal (the trend) out of the noise.

In contrast to a formal extrapolation, a forecast extrapolation is aimed at using various methods to seek out the simplest type of function which provides a maximum approximation to the trend of the process, considering its particular features and constraints and conforming to the hypotheses about its future development.

The procedure of selecting the type of approximating function in this instance, as a rule, includes a series of nonformal aspects such as assessing the conformity of the function to the points of a dynamic series. The very forecast research here consists of several stages usually of the following composition: the primary processing and reprocessing of the initial series; the choice of the type of extrapolation function; determining the parameters of the extrapolation function; the extrapolation itself and assessing the accuracy of the obtained results.

Let us point out that a formal extrapolation can be incorporated as a stage of research in a forecast extrapolation. On the other hand forecasts are frequently worked out on the basis of just formal extrapolation. In this regard let us examine it in more detail.

The mathematical basis for extrapolation and interpolation methods can be found in a section of function approximation in the theory of numerical analytical methods. The approximation problem is posed in the general case in the following manner [12]: a given function f(x) must be approximately replaced by a generalized polynomial

$$Q(x) = C_0\varphi_0(x) + C_1\varphi_1(x) + \cdots + C_m\varphi_m(x), \qquad (2.6)$$

so that the deviation, in a certain sense, of the function f(x) from Q(x) in the given set $X = \{x\}$ is the least. If the set X consists of a finite number of points $x_0, x_1, \ldots, x_n$, then the approximation is termed point, and if X is an interval $a \leqslant x \leqslant b$, then the approximation is termed integral. Let us examine the first instance.

The most important for practice are the degree polynomials of the type

$$Q(x) = a_0 + a_1x + a_2x^2 + \cdots + a_mx^m. \qquad (2.7)$$

In terms of them it is possible to formulate the approximation problem in the following manner: for the given function f(x) to find the polynomial Q(x) of the lowest possible degree m assuming at the set points $x_i (i = 1, 2, \ldots, n, x_i \neq x_j$ with $i \neq j)$ of the same value as the function f(x), that is, one where $Q(x_i) = f(x_i) (i = 1, 2, \ldots, n)$. The given system of points $x_1, x_2, \ldots, x_n$ is termed the basic point of interpolation while the polynomial Q(x) is the interpolation polynomial.

Let us point out that for extrapolation purposes it is possible to use the same interpolation polynomials in which are substituted the x values lying beyond the interval (a, b) or the given set of points X = {x}. However, mathematics recommends that extrapolation be very cautiously used. For example, in assigning the initial points in the form of a sequence of equidistant values with an interval h, it is recommended that orthogonal polynomes be used for an extrapolation only for a value $\Delta x = h/2$ or if the designated function has rather smooth ends.

On this level the forecaster must make much bolder conclusions, for example, in setting the initial series of values for a variable over a year to extrapolate this for 5 or even 10 years to come. Because of this there is the obvious necessity of involving certain additional information not found in the dynamic series.

However, let us return to the formal interpolation. If $n \leq m$, then it is possible to assume m = n and determine the factorization coefficients for a_i from the system of equations:

$$\left.\begin{array}{l} a_0 + a_1x_0 + \cdots + a_nx_0^n = y_0; \\ a_0 + a_1x_1 + \cdots + a_nx_1^n = y_1; \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_0 + a_1x_n + \cdots + a_nx_n^n = y_n. \end{array}\right\} \tag{2.8}$$

The determinant of this system is a Vandermond determinant $\Delta_{0 \leq p \leq q \leq n} = P(x_g - x_p) \neq 0$ and, consequently, the system (2.8) has a uniform solution. This solution will be called a Lagrangian interpolation polynomial and which can be written for the function f(x) in the explicit form

$$L_n(x) = \sum_{i=0}^{n} \frac{(x - x_0) \ldots (x - x_{i-1})(x - x_{i+1}) \ldots (x - x_n)}{(x_i - x_0) \ldots (x_i - x_{i-1})(x_i - x_{i+1}) \ldots (x_i - x_n)} y_i.$$

For the case of equidistant basic points of x_i, the polynomial $L_n(x)$ can be written in the form of a Newtonian interpolation polynomial.

If n > m, then a precise approximation is impossible, as the number of basic points exceeds the degree of the polynomial and the system (2.8) becomes incorrect. In this instance one moves over to an approximate interpolation, as a rule, to a point quadratic approximation. This comes down to seeking out the coefficients which minimize the quadratic deviation of the polynomial from the set points

$$S = \sum_{i=1}^{n} [Q_m(x_i) - f(x_i)]^2. \tag{2.9}$$

Such a method of approximation is called the least square method. If the degree of the approximating polynomial is greater than 3, then calculations using the least square method become extremely cumbersome and in addition the solution to systems of equations with high m values leads to significant losses in accuracy. Because of this a method is employed based on the use of orthogonal functions, that is, those where $\sum_{i=0}^{n} \phi(x_i)\psi(x_i) = 0$ in the entire set of points $X = \{x_0, x_1, \ldots, x_n\}$.

The approximating polynomials are constructed in a system of functions in which all functions of the system are paired orthogonal in the set $X = \{x_0, x_1, \ldots, x_n\}$.

Let $p_0(x), p_1(x), \ldots, p_m(x)$--the set system of orthogonal polynomials in the set $\{x_0, x_1, \ldots, x_n\}$, that is, $\sum_{i=0}^{n} p_j(x_i)p_k(x_i) = 0$, with $j \neq k$, and the polynomial indexes correspond to their degrees. In addition, let $S_j = \sum_{i=0}^{n} p_j(x_i) > 0$. Since the polynomials $p_j(x_i)(j = 0, 1, \ldots, m)$ are linearly independent, the arbitrary polynomial $Q_m(x)$ for the degree m can be represented in the form of their linear combination:

$$Q_m(x) = b_0p_0(x) + b_1p_1(x) + \ldots + b_mp_m(x).$$

In line with the orthogonality of the polynomials $p_j(x)$ it is possible to give in an explicit form a formula for calculating the coefficients for the factorization of b_j:

$$b_j = \frac{\sum_{i=0}^{n} Q_m(x_i)\, p_j(x_i)}{\sum_{i=0}^{n} p_j^2(x_i)} \quad (j = 0, 1, \ldots, m). \tag{2.10}$$

For using orthogonal polynomials in solving approximation problems the least square method is used (2.9). Let us examine a possible version for realizing a function approximation using the Chebyshev orthogonal polynomials.

Let us assume a general case when the function has been set by its values y_k in the set of points x_k. In this set, the Chebyshev orthogonal polynomials $p_0(x), p_1(x), p_2(x), \ldots, p_n(x)$ has been set so that $\sum_{k=1}^{N} p_i(x_k)p_j(x_k) = 0$ for all $i \neq j$.

Let us seek a representation of the given function in the form

$$Y = b_0p_0(x) + b_1p_1(x) + \ldots + b_np_n(x),$$

where the factorization coefficients are determined according to (2.10)

$$b_j = \frac{\sum_{\kappa=1}^{N} y_\kappa p_j(x_\kappa)}{\sum_{\kappa=1}^{N} p_j^2(x_\kappa)} \quad (j = 1, 2, \ldots, n). \tag{2.11}$$

The Chebyshev polynomials themselves have the following form with the leading term equal to 1:

$$p_0(x)=1;$$

$$p_1(x)=x-\bar{x};$$

$$p_2(x)=(x-\bar{x})^2-\frac{\sum_{\kappa=1}^{N}(x_\kappa-\bar{x})^3}{\sum_{\kappa=1}^{N}(x_\kappa-\bar{x})^2}(x-\bar{x})-$$

$$-\frac{\sum_{\kappa=1}^{N}(x_\kappa-x)}{N}=x^2-\frac{\overline{x^3}-\overline{x^2}\bar{x}}{\overline{x^2}-(\bar{x})^2}\,x+\frac{\overline{x^2}x-(\bar{x})^2}{\overline{x^2}-(\bar{x})^2}\text{ etc.} \qquad (2.12)$$

In organizing the calculation process for figuring the Chebyshev polynomials, a recurrent procedure is employed and this is described by the following formulas:

$$p_{j+1}(x)=(x+\beta_{j+1})\,p_j(x)-\frac{H_j}{H_{j-1}}\,p_{j-1}(x)\ (j=1,2,\ldots,n);$$

$$H_j=\sum_{\kappa=1}^{N}p_j^2(x_\kappa)\,\omega_\kappa;\ \beta_{j+1}=-\frac{1}{H_j}\sum_{\kappa=1}^{N}x_\kappa p_j^2(x_\kappa)\,\omega_\kappa$$

$$(j=0,1,2,\ldots,n).$$

Fig. 2.6 shows the block diagram of an algorithm for the calculation process in figuring the coefficients of Chebyshev polynomials following the above-described procedure. It uses the following designations: X--the vector for the values of the independent variable representing the set $\{x_0, x_1, \ldots, x_N\}$; Y--the vector for the values for the function representing the set $\{y_0, y_1, \ldots, y_N\}$; N--the total number of set points; n--the maximum degree of the polynomial; B--the vector for the values of the factorization coefficients $\{b_j\}$; p0[k]--corresponds to $p_{j-1}(x_k)$; p1[k] corresponds to $p_j(x_k)$; pa, pb, pc--working values corresponding to $p_{j-1}(x)$, $p_j(x)$, $p_{j+1}(x)$ in the calculation process; ha, hb correspond to H_{j-1} and H_j; BETA--corresponds to β_{j+1}.

In the block diagram there is the subprogram PP-"B" for calculating the values b_j, H_{j-1}, H_j, β_{j+1} using the formulas (2.11), (2.13) and (2.14).

After calculating the coefficients we obtain an analytical representation of the point aggregate. However, as an examination of the algorithm indicates, it is not simple to represent its notation in the form of a degree sequence with coefficients because of the recurrent calculations of the polynomials themselves. For utilizing the obtained approximation in practice it is essential to employ an electronic computer as a means of storing the information and figuring the values of the y function for any values of the independent variable x. In the computer storage it is enough to have the values for the coefficients B = $\{b_j\}$; BETA = $\{\beta_j\}$; H = $\{H_j\}$; $\bar{x}$. Then using formula (2.13) it is possible to calculate in sequence any polynomial starting from $p_2(x)$ and $p_0(x) = 1$; $p_1(x) = x - \bar{x}$. Then, having multiplied the

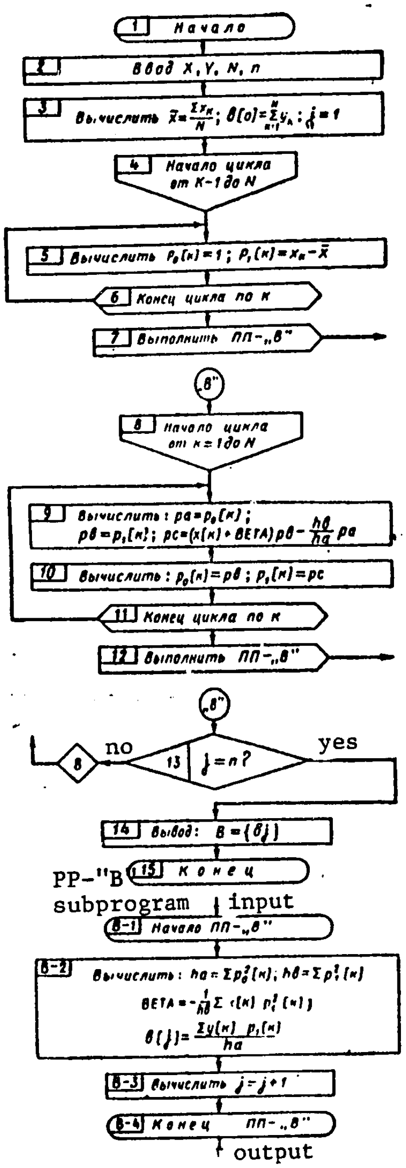


Fig. 2.6. Block diagram of algorithm for calculating coefficients of Chebyshev interpolation polynomial

Key: [only Russian translated in the appropriate box]
1--Start; 2--Input; 3--Calculate; 4--Start of cycle from k-1 to N; 5--Calculate; 6--End of cycle for k; 7--Carry out PP [subprogram] - "B"; 8--Start of cycle from k = 1 to N; 9--Calculate; 10--Calculate; 11--End of cycle for k; 12--Carry out PP-"B"; 14--Conclusion; 15--End; B-1 start of PP-"B"; B-2 calculate; B-3 calculate; B-4 end.

obtained values by the b_j coefficients, it is possible to obtain the values of the function y = f(x) with the accepted degree of approximation.

Very widefound is the instance when the values of the independent variable x are taken with an even interval $h = x_{k+1} - x_k (k = 1, 2, \ldots, N-1)$. In this instance the calculating of Chebyshev polynomials is significantly simplified.

Let us substitute the variables $u = (x-\bar{x})/h$, where $\bar{x} = (x_1+x_N)/2$. The sought polynomial is written in the form

$$y = c_0 p_0(u) + c_1 p_1(u) + \ldots + c_n p_n(u); \qquad (2.15)$$

$$c_j = \frac{1}{H_j}\sum_{k=1}^{N} y_k p_j(u); \quad (j = 0,\ 1,\ \ldots,\ n;\ n < N),$$

where $H_j = \sum\limits_{k=1}^{N} p_j^2(u_k)$; $u_k = (x_k-x)/h$; $p_0(u) = 1$; $p_1(u) = u$, and the recurrent formula

has the form $p_{j+1}(u) = up_j(u) - \frac{H_j}{H_{j-1}} p_{j-1}(u)$ $(j = 1,\ 2,\ \ldots,\ n)$.

The values H_j and the ratios H_j/H_{j-1} in these formulas are calculated through the total number of points N:

$$H_0 = N; \quad H_1 = \frac{N(N^2-1)}{12}; \quad H_2 = \frac{N(N^2-1)(N^2-4)}{180} \text{ and so forth,}$$

and there are tables for H_j for the various N and this facilitates the calculation procedure.

There is a significant number of different polynomials making it possible to carry out interpolation and extrapolation using various approximation formulas. These are the formulas of Lagrange, Newton, Stirling, Legendre, Lagger and others. Very widespread are the methods of a harmonious analysis of periodic processes using trigonometric polynomials of the type

$$y = \frac{a_0}{2} + \sum_{j=1}^{m} (a_j \cos j\omega x + b_j \sin j\omega x); \quad \omega = \frac{2\pi}{T}, \qquad (2.16)$$

where a and b--polynomial parameters;
T--the period for the change of function.

With equal accuracy of the measurements and equidistant values for the independent

variable $x_\kappa = \frac{T}{N}\kappa$ $(\kappa = 0,\ 1,\ \ldots,\ N-1)$:

$$a_j = \frac{2}{N}\sum_{\kappa=0}^{N} y_\kappa \cos j\omega x_\kappa \quad (j = 0,\ 1,\ \ldots,\ m); \qquad (2.17)$$

$$b_j = \frac{2}{N}\sum_{\kappa=0}^{N-1} y_\kappa \sin j\omega x_\kappa \quad (j = 0,\ 1,\ \ldots,\ m),$$

where $x = \frac{2\pi}{N}k$.

It is convenient to construct the calculation procedure for an uneven number of points N = 2L + 1 using the reccursive procedure. The sequence is calculated recurrently:

$$u_{2L,j} = y_{2L}; \quad u_{2L-1,j} = y_{2L-1} + 2\cos\omega x_j y_{2L} \tag{2.18}$$

and hence

$$u_{\kappa,j} = y_\kappa + 2\cos\omega x_j u_{\kappa+1,j} - u_{\kappa+2,j}$$
$$(\kappa = 2L-2,\ 2L-3,\ \ldots,\ 2,\ 1),$$

then

$$a_j = \frac{2}{N}(y_0 + u_{1,j}\cos\omega x_j - u_{2,j}); \tag{2.19}$$

$$b_j = \frac{2}{N} u_{1,j}\sin\omega x_j.$$

Thus, it is possible to avoid repeated calculations of the sin and cos for the square values of the independent variable. This substantially reduces the calculation. The procedure's program is rather simple.

Very essential in problems of this sort is the question of selecting the amount of the degree of the polynomial in the factorization. The optimum degree of the polynomial is determined by the following considerations. A low degree polynomial gives a rough approximate description of the process and with a high degree of the polynomial the effect of smoothing and the filtering of random deviations (noise) is lost and this obscures the essence of the studied phenomenon. On the other hand with a rise in the degree of the approximating polynomial there is a sharp rise in the machine time expenditures for the calculations as well as the memory for storing the information.

Let a polynomial of the degree n_0 represent absolutely accurately the functional dependence set by the points y = f(x), and the measured values y_k contain the random errors η_k distributed according to a normal law. They are independent and have equal variance σ^2 [31]. Let the factorization be carried out according to Chebyshev orthogonal polynomials (2.11)-(2.14).

The random errors η_k cause mistakes in the factorization coefficients b_j (let us designate them β_j) and give rise to deviations in the changed values of the function from the calculated. If we designate the coefficients obtained on the basis of an experiment by b_{je}, then the deviations at this point are determined as follows:

$$R_\kappa = y_\kappa - \sum_{j=0}^{n_0} b_{je} p_j(x_\kappa). \tag{2.20}$$

In the notation through the measurement errors

$$R_\kappa = \eta_\kappa - \sum_{j=0}^{n_0} \beta_j p_j(x_\kappa). \tag{2.21}$$

The deviations will follow a normal law with the variance

$$\sigma^2_{\text{отк}} = \sigma^2 \left(\frac{1}{\omega_\text{к}} - \sum_{j=0}^{n_0} \frac{1}{H_j} p_j^2 (x_\text{к}) \right). \tag{2.22}$$

The minimum total of the deviation squares

$$S_{n_0} = \sum \omega_\text{к} R_\text{к}^2 = \sum_{\text{к}=0}^{N} \omega_\text{к} \left[y_\text{к} - \sum_{j=0}^{n_0} b_{j0} p_j (x_\text{к}) \right]^2 \tag{2.23}$$

has an average value

$$MS_{n_0} = \sigma^2 (N - n_0 - 1). \tag{2.24}$$

Here ω_k--the weight multipliers with different measurements of $y_k(x_k)$.

The ratio

$$S_{n0}/(N - n_0 - 1) \tag{2.25}$$

serves as an unbiased consistent estimate for the variance σ^2. With a sufficiently large $N - n_0 - 1$, it is possible to consider $\sigma^2 \approx S_{n0}/(N - n_0 - 1)$.

This ratio can serve as an unbiased consistent estimate of the variance with any $n \geqslant n_0$.

The procedure for selecting the optimal degree of a polynomial can be represented in the following form:

1) Let us calculate the factorization coefficients b_0, b_1, b_2 according to the algorithm given in Fig. 2.6;

2) For each next value of b_j let us figure the total of the squares of deviations S_n (2.23) which can be more conveniently figured using the formula

$$S_n = \sum_{\text{к}=1}^{N} \omega_\text{к} y_\text{к}^2 - (b_0^2 H_0 + b_1^2 H_1 + \cdots + b_n^2 H_n);$$

The addition of each new term in the factorization of y for the orthogonal polynomials reduces S_n by the amount $b_{n+1}^2 H_{n+1}$;

3) Each calculated next S_n is divided by the number $N - n_0 - 1$ and the obtained deviation of the type (2.25) is compared with its previous $(n - 1)$ value;

4) The process terminates when the ratio (2.25) ceases to decline noticeably; this last value of n is also taken as the optimum degree of the polynomial n_0; the threshold of the diminishing rate can be set differently, for example, by the condition

$|S_{n-1} - S_n| \leqslant \varepsilon$ or by the relative value $S_{n-1}/S_n < 1+\delta$, where ε and δ --the set small positive amounts.

The algorithm is sufficiently simple and can be directly incorporated in the block diagram for the algorithm of figuring the factorization coefficients for the Chebyshev polynomials (see Fig. 2.6). Analogous estimates and rules can be provided for selecting the optimum procedure for a trigonometric polynomial.

Thus, we have examined the basic questions related to seeking out an approximate description of a function in an interval using polynomials. This method is the most typical in solving interpolation and extrapolation problems by formal mathematics. Here we have not discussed the questions of selecting the most suitable type of orthogonal polynomials for one or another numerical sequence. Sufficiently universal procedures for solving this problem as yet have not been worked out although the first steps have been taken in the direction of formalizing the procedure of choosing the type of functions which ensure an optimum approximation [19, 20].

2.4.2. The Use of the Method of Finite Differences in Solving Series Extrapolation and Interpolation Problems

In setting a numerical series with a constant interval in the increment of the independent variable in solving interpolation and extrapolation problems, often instead of the very values of the function it is advisable to employ series of its sequential increments. Such a procedure in a number of instances is very effective in solving forecasting problems and is presently rather widely used in the practice of various modifications of extrapolation forecasts.

Let us examine the essence of this method. The formation of finite differences with a reduction in the increase interval of the independent variable leads to an operation which is close to a repeated differentiation of a continuous function, as to a reduction in the order of the curve and to a simplification of its analysis and extrapolation.

Let the function be set by a numerical sequence of values with equidistant spacing of h: x_0, y_0; x_0+h, y_1; ...; x_0+ih, y_i; ...; x_0+nh, y_n. The amount $y_{i+1}-y_i$ is termed the first descending difference and is designated Δy_i, the amount $\Delta y_{i+1} - \Delta y_i$ is termed the second descending difference $\Delta^2 y_i$ and so forth, $\Delta^n y_i = \Delta^{n-1} y_{i+1} - \Delta^{n-1} y_i$. It is convenient to reduce the values of the descending differences to a table of the sort given below.

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
x_0	y_0	Δy_0	$\Delta^2 y_0$	$\Delta^3 y_0$
x_1	y_1	Δy_1	$\Delta^2 y_1$	
x_2	y_2	Δy_2		
x_3	y_3			

In the simplest case we proceed in the following manner. In sequence the finite differences are calculated for the first, second and so forth orders for the initial numerical series with a set interval for the increment of the independent variable. The order of the differences increases until the difference of order m becomes virtually constant within the limits of a certain research accuracy. If the number of points of the initial series n is slight, then it may seem that stabilization does not occur up to the last $\Delta^n y$ difference. In this case, ordinarily the n order is considered the one at which there has been a stabilization of the differences in value. Thus we have:

initial series $y_0 = f(x_0)$, $y_1 = f(x_0 + h)$, $y_2 = f(x_0 + 2h)$, ..., $y_n = f(x_0 + nh)$;

first order differences $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, ..., $\Delta y_{n-1} = y_n - y_{n-1}$;

second order differences $\Delta^2 y_0 = \Delta y_1 - \Delta y_0$, $\Delta^2 y_1 = \Delta y_2 - \Delta y_1$, ..., $\Delta^2 y_{n-2} = \Delta y_{n-1} - \Delta y_{n-2}$;

m order differences $\Delta^m y_0 = \Delta^{m-1} y_1 - \Delta^{m-1} y_0$, $\Delta^m y_1 = \Delta^{m-1} y_2 - \Delta^{m-1} y_1$, ..., $\Delta^m y_{n-m} = \Delta^{m-1} y_{n-m+1} - \Delta^{m-1} y_{n-m}$.

Let us assume that $\Delta^m y_0 \approx \Delta^m y_1 \approx \ldots \approx \Delta^m y_{n-m} \approx \Delta^m y \approx$ constant. Then let us accept that $\Delta^m y_{n-m+1} = \Delta^m y$, that is, that the finite m-order differences maintain their constancy for the following point of the series. Let us carry out the reverse calculation:

$$\Delta^{m-1} y_{n-m+2} = \Delta^{m-1} y_{n-m+1} + \Delta^m y_{n-m+1},\ \Delta^{m-2} y_{n-m+3} =$$
$$= \Delta^{m-2} y_{n-m+2} + \Delta^{m-1} y_{n-m+2},\ \ldots,\ \Delta y_n = \Delta y_{n-1} + \Delta^2 y_{n-1},$$
$$y_{n+1} = y_n + \Delta y_n.$$

Thus, we have obtained the point $y_{n+1} = f\{x_0 + (n+1)h\}$, that is, the value of the function for the next increase step of the independent variable beyond the limits of the initial series of numerical values.

Having assumed furthermore the constancy of difference m for the following point $\Delta^m y_{n-m+2} = \Delta^m y$ and carrying out an analogous calculation, we obtain the value of the function on the following increment interval for the independent variable and so forth.

This method is very approximate and leads to satisfactory results with low degrees in the factorization of a function, with a smooth type of initial series (this is the main thing) and for a small number of extrapolation intervals.

Special Newtonian formulas are used for the extrapolation and interpolation of functions using the method of finite differences. These have been also used for standard programs to organize the corresponding computer calculations. The Newtonian formula for an interpolation forward has the form

$$P_n(x) = y_0 + t\,\Delta y_0 + \frac{t(t+1)}{2!}\Delta^2 y_0 + \cdots + \frac{t(t-1)\ldots(t-n+1)}{n!}\Delta^n y_0, \qquad (2.26)$$

where t--the number of intervals between x and x_0, $t = (x-x_0)/h$;
$\Delta^k y_0$--the finite difference of order k.

If we assume n = 1, then we obtain a formula for a linear interpolation

$$P_1(x) = y_0 + t\Delta y_0.$$

With n = 2, we obtain a quadratic interpolation

$$P_2(x) = y_0 + t\Delta y_0 + \frac{t(t-1)}{21} \Delta^2 y_0.$$

If the values of the functions are rather numerous, then the number n is chosen in such a manner that the difference $\Delta^n y_i$ be constant with the set degree of accuracy with the limited number of points n = m-1, where m--the total number of set points.

The error of interpolation for (2.26) is determined in the following manner:

$$R = \frac{f^{(n+1)}(\xi)}{(n+1)!} t(t-1)\ldots(t-n)h^{(n+1)}, \tag{2.27}$$

where $x_0 \leqslant \xi \leqslant x$. The error ξ will be minimal with small values of t, that is, in the vicinity of the initial value of x_0.

There is an analogous Newtonian formula for an interpolation backwards:

$$P_n(x) = y_n + t\Delta y_{n-1} + t\frac{t+1}{2!}\Delta^2 y_{n-2} + \cdots + \frac{t(t+1)\ldots(t+n-1)}{n!}\Delta^n y_0. \tag{2.28}$$

It is advisable to employ formula (2.28) for $x > x_n$. Formulas (2.26) and (2.28) can be used for extrapolation of the y values. For $x \leqslant x_0$ it is advisable to use (2.26), and $t = (x-x_0)/h < 0$. With $x > x_k$ and x close to x_k, it is convenient to use (2.28) and $t = (x-x_0)/h > 0$.

For calculating one value of the function f(x) using any of these formulas, there must be n divisions, $\left(\frac{n^2+3n}{2} - 1\right)$ multiplications and (n^2+1) additions and altogether $(1.65n^2 + 4.85n - 0.3)$ conditional arithmetic operations. In knowing the speed of a computer it is possible to assess the machine time expenditures for calculating one point on the standard interpolation (extrapolation) program.

2.4.3. The Methods of Preliminary Processing and Presenting Initial Data in a Forecast Extrapolation

In contrast to a formal extrapolation, a forecast extrapolation does not come down merely to calculating dependence factorization coefficients for apriori selected polynomials or to calculating the parameters of a predetermined function. In a forecast extrapolation, in the process of analyzing the set series of values as well as the essence of the described process, hypotheses are advanced on the nature of its further development and on the basis of them the type of extrapolating function is chosen.

As experience of working out such forecast models shows, very essential here is the method of presenting the initial data and the procedure of their preliminary processing. Let us examine certain possibilities for realizing these procedures.

With a large number of empirical points, the choice of the extrapolation function can be facilitated by smoothing the initial series. As was said above, for a forecast extrapolation the essential thing was to eliminate the random deviations (noise) from the experimental sequences.

Smoothing is carried out by polynomials which approximate the groups of experimental points using the least square method. The best smoothing is obtained for the midpoints of the group and for this reason it is desirable to choose an uneven number of points in the group to be smoothed. The very groups of points are taken by composition as moving through the entire table. For example, for the first five points y_1, y_2, y_3, y_4, y_5 the average y_3 is smoothed and then for the following five y_2, y_3, y_4, y_5, y_6 the y_4 is smoothed and so forth.

The remaining end points are smoothed using special formulas. The most widespread form of smoothing is linear, that is, using a first-degree polynomial. For smoothing for three points, the formulas are as follows:

$$\begin{aligned} \tilde{y}_0 &= \frac{1}{3}(y_{-1} + y_0 + y_1); \\ \tilde{y}_{-1} &= \frac{1}{6}(5y_{-1} + 2y_0 - y_1); \\ \tilde{y}_{+1} &= \frac{1}{6}(-y_{-1} + 2y_0 + 5y_1), \end{aligned} \tag{2.29}$$

where y_0, $\tilde{y}_0$--values of the initial and smoothed function at the midpoint;
y_{-1}, $\tilde{y}_{-1}$--values of initial and smoothed function to the left of the midpoint;
y_{+1}, $\tilde{y}_{+1}$--values of the initial and smoothed function to the right of the midpoint.

The formulas for $\tilde{y}_{-1}$ and $\tilde{y}_{+1}$ are used, as a rule, only for the ends of the interval.

Analogous formulas exist for smoothing series for five points:

$$\begin{aligned} \tilde{y}_0 &= \frac{1}{5}(y_{-2} + y_{-1} + y_0 + y_{+1} + y_{+2}); \\ \tilde{y}_{-1} &= \frac{1}{10}(4y_{-2} + 3y_{-1} + 2y_0 + y_1); \\ \tilde{y}_{+1} &= \frac{1}{10}(y_{-1} + 2y_0 + 3y_1 + 4y_2); \\ \tilde{y}_{-2} &= \frac{1}{5}(3y_{-2} + 2y_{-1} + y_0 - y_2); \\ \tilde{y}_{+2} &= \frac{1}{5}(-y_{-2} + y_0 + 2y_1 + 3y_2). \end{aligned} \tag{2.30}$$

A block diagram for an algorithm to smooth a sequence of points according to the formulas of (2.30) for a group of five points can be represented in the form given in Fig. 2.7.

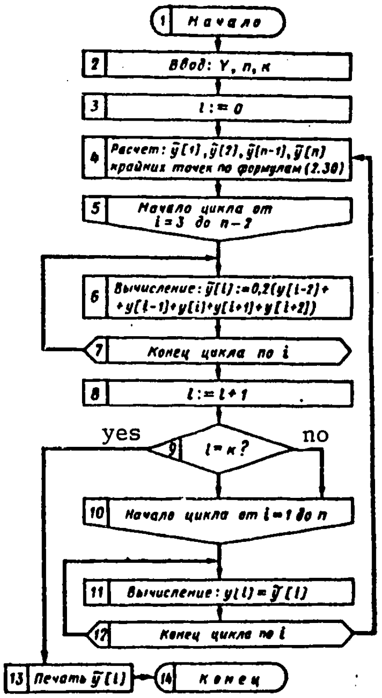


Fig. 2.7. Block diagram of algorithm for smoothing a sequence of n points

Key: [only Russian terms translated in appropriate boxes] 1--Start; 2--Input; 4--Calculation... of end points from formulas (2.30); 5--Start of cycle from $i = 3$ to $n = 2$; 6--Calculation; 7--End of cycle for i; 10--Start of cycle from $i = 1$ to n; 11--Calculation; 12--End of cycle for i; 13--Print; 14--End

Smoothing, even in a simple linear version, is in many instances a very effective method for disclosing a trend in superimposing random interference and measurement errors on an empirical numerical series. In the block diagram shown in Fig. 2.7 a possibility has been provided for carrying out repeated smoothing of the initial numerical series. The number of sequential smoothing cycles is set by the value K.

The choice of the amount of K should be made depending upon the type of initial series, upon the degree of its assumed distortion by noise and upon the goal pursued by the smoothing. Here it is essential to bear in mind that the effectiveness of this procedure declines rapidly (in a majority of instances) so, as experience shows, it is advisable to repeat it from one to three times.

As a certain objective criterion from which it is possible to judge the inadvisability of a repeated smoothing, one can use the expression: $\max\{|y_i - y_j\}| \leqslant \varepsilon$, where ε--a positive number chosen from considerations of the accuracy of data presentation and the accuracy of the subsequent processing algorithms.

An example of the processing of a numerical series using the smoothing method is illustrated in Fig. 2.8. The broken line shows the initial dynamics for passenger departures over the years for one of the nation's airfields. Curves 1 and 2 correspond to the smoothed series for three and five points.

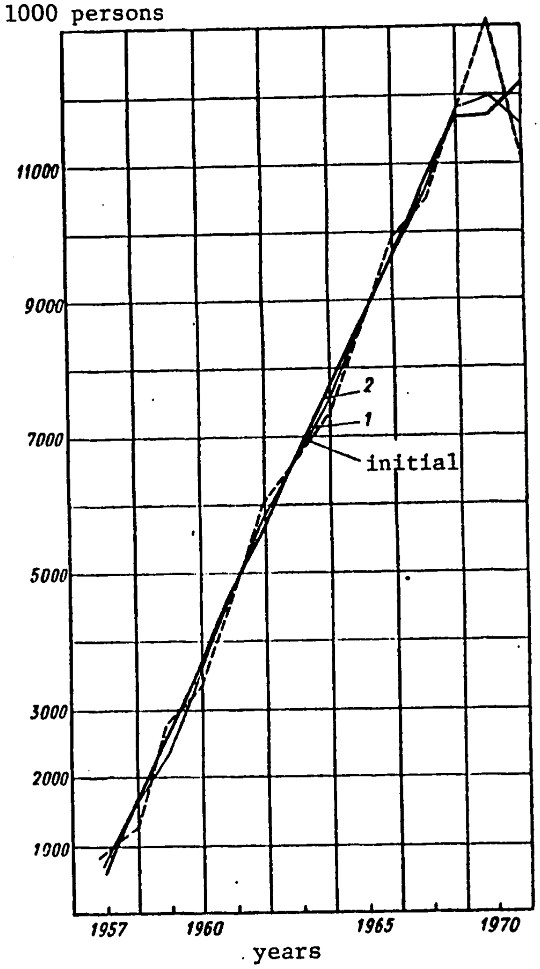


Fig. 2.8. Processing numerical series using smoothing method

In addition to the linear smoothing, a significant number of formulas are known for nonlinear smoothing by higher-degree polynomials. However, in practice they are employed comparatively rarely (at least above the third degree) for the reason that for a satisfactory realization they require only large-sized table, in addition, the edges of the table are not sufficiently well smoothed and the formulas themselves become cumbersome. Nevertheless, in the case of large initial files, complex types of curves and the use of computers, their application is fully justified.

While smoothing is a procedure aimed at the primary processing of a numerical series for the purpose of excluding random fluctuations and elucidating a trend, fitting is used for the more convenient presentation of the initial series, in leaving the numerical values as before. Fitting is the name given to the reduction of the empirical formula y = f(x, a, b) to the type

$$Y = a_1X + b_1. \qquad (2.31)$$

The formula (2.31) examines a two-parameter initial function. This is due to the fact that this function is the most widely found in the practice of extrapolation and interpolation calculations, and on the other hand, in a majority of instances is comparatively easy to fit. Functions with a greater number of parameters are far more difficult to fit and in fact cannot always be done so.

The most common fitting procedures are taking logorithms and the change of variables. Let us examine these procedures from a series of the following specific examples.

1. For finding the parameters of the exponential function $y = ax^b$, a logorithmic transformation is employed of the type: $\lg y = \lg a + b \lg x$ and the change of variables $X = \lg x$, $Y = \lg y$. As a result we have (2.31), where $a_1 = b$, $b_1 = \lg a$.

Thus, having rearranged the experimental points of the proposed exponential dependence in the logorithmic ruling, we obtain a linear dependence which can easily be described and extrapolated and then recalculate the results using the formulas inverse to the initial transformation of the variables.

2. For the exponential function $y = ae^{bx}$ it is also possible to apply logorithmic fitting: $\lg y - \lg a + b \lg e\ x$ and the change $X - x$, $Y = \lg y$. We obtain (2.31), where $a_1 = b \lg e$, $b_1 = \lg a$. In this instance it is essential to provide for the rearrangement of the exponential points in a semilogorithmic scale with the subsequent analysis of the obtained graph.

3. For dependences of the type: a) $y = 1/(ax+b)$ and b) $y = x/(ax+b)$, the following transformations are used: a) $Y = 1/y = ax+b$, b) $X = 1/x$ and $Y = 1/y$. This gives $Y = \left(\frac{a}{X} + b\right) \Big/ \frac{1}{X} = a+bX$. In this instance along the axes of the grid one must lay off the amounts inverse to the values of the initial variables.

4. If the proposed empirical dependence has the form $y = 1/(a+be^{x})$, then the transformation of fitting is $Y = 1/y$, $X = e^{-x}$. Then the coefficients of (2.31) will be $a_1 = b$, $b_1 = a$.

It is essential to bear in mind that the values of the function parameters determined after the fitting minimize the total squares of deviations for the transformed values from the linear dependence of (2.31) and do not always correspond to the minimum deviation of the measured values from the calculated. For this reason such a calculation must be considered only a certain approximation to the truly optimum coefficient values.

In the case that the empirical formula is assumed to contain three parameters or it is known that the function is a three-parameter one, then by certain transformations it is sometimes possible to exclude one of the parameters and the remaining two can be reduced to one of the fitting formulas.

For example, the initial formula $y = ax^{b}+c$ can be fitted after an approximate calculation of the c parameter. For this we select x_1, x_2 which are moved farther apart in the empirical series and x_3 which is linked to them by the ratio $x_3:x_1 = x_3:x_2$. For such a choice of independent variables one approximately determines $c = (y_1y_3 - y_2^3)/(y_1+y_3-2y_2)$. Then by the changing of variables $X = \lg x$; $Y - \lg (y-c)$, the initial formula is reduced to a linear one: $Y = bX + \lg a$. In comparing it with (2.31) we have $a_1 = b$ and $b_1 = \lg a$.

After defining the parameters it is recommended that for all the points the correctness of the approximate calculation of c be tested and in the instance of significant discrepancies recalculate the parameters.

It is possible to view fitting not as a method of presenting initial data but rather as a method for a direct approximate determination of parameters for a function which approximates the initial numerical series. This method is often employed precisely in this manner in certain extrapolation forecasts. We would point out that the possibility of its direct use for defining the parameters of an approximating function is determined chiefly by the type of the initial numerical series and by the degree of our knowledge and by our confidence about the type of function describing the studied process.

In the event that the type of function is unknown, fitting must be viewed as a preliminary procedure in the process of which, by employing various formulas and procedures, the most suitable type of function describing the empirical series is ascertained.

Graphic and numerical analysis of the dependence of two variables could be the sole method for selecting the form of connection if a set of equally correct analytical expressions did not conform to the same function graph or ratio. Fig. 2.9 gives several examples of analytical expressions where it would be impossible to choose one in using the above-given methods of preliminary analysis.

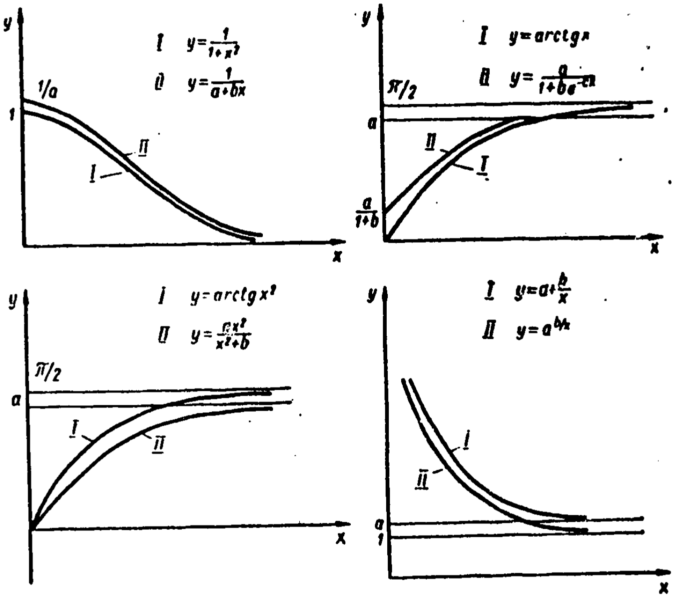


Fig. 2.9. Graphic depiction of various analytical expressions

The examination of an empirical series for the purpose of elucidating the optimum type of function describing it is a broadening of the fitting method or its generalization. Here it is not essential that the transformations lead to linear forms. However their results prepare and facilitate the process of choosing the approximating function in the problems of forecasting extrapolation.

Source [43] gives a procedure for such research using differential growth functions. In the simplest case it is proposed that the following three types of differential growth functions be employed.

1. The first derivative or absolute differential growth function

$$\phi(t) = y' = \frac{dy}{dt}. \tag{2.32}$$

On the graph $\phi = f(t)$ this is represented by an angular coefficient at each point of the graph. The $\phi(t)$ = const for the linear law of change of $y(t)$. For second order curves (parabolic laws) $\phi(t)$ has a linear type of change and for the exponential curves of $\phi(t)$ also an exponent. The value of $\phi(t)$ depends upon the selected scale for the measuring of the exponent and time.

2. The relative differential coefficient or the logorithmic derivative

$$\omega(t) = \frac{dy}{dt} \Big/ y = d(\log y)/dt. \tag{2.33}$$

This function can be shown on a graph by constructing it on a semilogorithmic scale. Then $\omega(t)$ will be an angular coefficient at each point. For the exponential dependents $\omega(t)$ = const and for the exponential function $\omega(t)$ there is a hyperbolic nature.

3. Elasticity of the function $\varepsilon(t) = \frac{t\,dy}{y\,dt} = \frac{d(\log y)}{d(\log t)}$. (2.34)

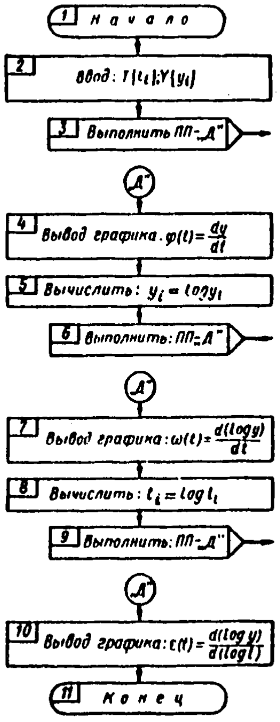


Fig. 2.10. Block diagram for calculating differential growth functions

Key: [translation of only Russian in appropriate block]

1--Start; 2--Input; 3--Carry out PP [subprogram]-"D"; 4--Derivation of graph; 5--Calculate; 6--Carry out PP-"D"; 7--Derivation of graph; 8--Calculate; 9--Carry out PP-"D"; 10--Derivation of graph; 11--End

In a graph of a dynamic series constructed on a logorithmic scale, elasticity is defined as an angular coefficient at each point. The $\varepsilon(t)$ = const for an exponential function; for the exponential function $\varepsilon(t)$ has a linear type of change and it is also linear for the combined exponential function.

It must be pointed out that the elasticity of $\varepsilon(t)$ is a dimensionless amount and this makes it possible to employ it in comparing the nature of changes in different processes occurring in the most different possible time scales.

Source [43] gives graphs for differential growth functions for all the approximating functions most used in forecasting extrapolation, including: linear, parabolic, exponential, logistical, hyperbolic and so forth. An examination of growth functions indicates that by their combining it is possible rather uniformly to define the type of function producing them and determined by a numerical empirical series.

Thus, for preparing to take a decision on the type of approximating function, it is possible to propose a machine procedure for calculating the differential growth functions following the scheme shown in Fig. 2.10. It is best to carry this out on computers having a graphic data output (graph plotter, teletype, electric typewriter or display). The set of graphs obtained on the output for $\phi(t)$, $\omega(t)$, $\varepsilon(t)$ is compared with a standard reference table of the type given in [43]. In the block diagram of Fig. 2.10, the PP-"D" designates a subprogram for differentiating the function of $y(t)$ set by the series $\{y_i, t_i\}$.

2.4.4. The Problem of Choosing the Type of Function and the Ways of Sovling It for a Forecast Extrapolation

In the process of smoothing a dynamic series, in fitting it and determining the functions of differential growth, the type of function describing the initial process is already approximately determined and sometimes the estimates of this function's parameters are even obtained. For the final choice of the type of function a retrospective series study carried out in the stage of preliminary processing must be supplemented by research on the logic of the occurrence of the process as a whole, including hypotheses on its occurrence in the future and by research on the physical essence of the process, possible shifts, jumps and constraints stemming from this essence.

The basic questions which the researcher should set for himself at this stage are: 1) is the studied indicator as a whole an amount which grows uniformly, diminishes uniformly, is stable or has an extremum (or several of them) or is periodic; 2) is the studied indicator limited above (or below) by any limit; 3) does the function defining the process have a bending point; 4) does the function representing the process possess the property of symmetricalness; 5) does the process have a clear limitation of development in time.

Depending upon the answers to each of the listed basic questions, secondary questions arise which are of a more quantitative nature or the nature of elucidating the reasons for the appearance of one or another quality in the process.

For the answer to the first question it is essential to bring together information obtained in the process of the primary processing of the series and namely the first derivative $\phi(t)$ and the general considerations on the nature of the process's development. Obviously for a steadily rising function the graph $\phi(t)$ should lie completely in the positive area and its extrapolation would not show a tendency to cross the abscissa axis in the future.

Now about the essence of the process itself. For example, a majority of parameters determined by the development of scientific and technical progress can be steadily growing functions which in a number of instances have asymptotic constraints. Thus it is possible to speak about the continuous increase in the speed of transport, the rate of data processing, the power of energy units, the distance man penetrates into space, the increase in the length of human life, the greater labor productivity and so forth.

Analogous arguments can be given for steadily diminishing processes. For example, the shortening of production cycles, the reduction of relative dimensions and weight of units and so forth.

For stable parameters it is possible to estimate the amount of their degree of instability in the retrospective interval Δy_{max} or $\delta_{max} = \Delta y_{max}/\bar{y}$. It is essential to bring out the factors which influence the instability and analyze their possible changes in the future.

The most complicated problem in forecasting is the predicting of the appearance of abrupt jumps in the studied process in the future. The basic ways for solving this in the area of scientific-technical and economic forecasting are research using lead methods on patent and scientific-technical information as well as carefully planned expert surveys.

The extremums on the retrospective interval are easily detected in examining the graph of $\phi(t)$ at the points it crosses the abscissa axis. The presence of extremums in the previous development of the process leads to questions about their causes and their possible occurrence in the future. Here it is essential to examine the stability of their appearance in the development process and draw a conclusion on the possible periodic nature of the process.

In the event of a supposition of the monoextremalness of the process, the basic question is to elucidate the point of the extremum and the extremal value of the forecasted parameter (if the extremum has not yet been passed). This problem can be solved by the extrapolation of $\phi(t)$ and by determining the point of its crossing the time axis. This research must be supplemented by the results of qualitative and quantitative analysis for the development of factors influencing the achieving of an extremum by the examined process.

Many economic development processes are characterized by a periodic nature of development. In truth, these processes basically have a cyclical and not purely a periodic nature. Thus, the cyclical nature of economic development under capitalism is generally known. The process of its development contains a number of repeating stages: economic increase, crisis, depression, a new increase and so forth. Econometric research has shown that over long intervals of time the period of the cycle can remain relatively stable, although the height of the rise and the depth of the fall can change from cycle to cycle within certain limits.

If in the research process it is discovered that a certain indicator depends substantially upon the cyclical development of the overall economic process, then this fact provides essential information for selecting the type of function representing it in the extrapolation research. In particular, the methods of harmonic analysis and factorization for trigonometric polynomials (2.16)-(2.19) are used very effectively for such processes.

The reply to the second basic question about the nature of the process is very essential for selecting the correct type of function which extrapolates the trend. We have already mentioned at the beginning of (2.4) the mistakes to which extrapolation can lead in forecasting with a failure to consider the possible constraints of the process. For example, a dynamic series describing an increase in the speed of transport over the last 100 years is described rather well by the exponential dependence [46]. However, consideration of the natural limit of the speed of light inclines one to choose an S-shaped curve (logistical) for describing the process. The forecast results using these two curves differ substantially.

In any area of knowledge great attention is given to the problem of studying development limits. In the area of scientific and technical forecasting it is possible to point out several types of limits examined below.

Absolute limits are unconditional limits where the area of action is not restricted. Among absolute limits are [50], for example: the speed of light, absolute temperature zero, zero pressure, an efficiency of 1, the temperature for the breaking of molecular bonds and so forth.

Relative limits occur in a certain area or in terms of a certain object. These include the terrestrial limits such as: maximum speed in the atmosphere, maximum depth of the ocean, minimum speed for the orbiting of a sattelite and so forth. It is also possible to give examples of the relative limits of human capabilities: maximum g-loads, maximum noise level and so forth.

Calculated limits are more particular limits set on the basis of the first two types of limits and various sorts of transformations and laws linking them to the derived amounts. For example, the maximum value of efficiency with a set temperature drop in the Carnot cycle, the limit for loading storage with information of 10^{14} bits per cm^3 calculated on the basis of the absolute limit determined by the Heisenberg formula; the maximum amount of microminiaturization of electronic elements related to the Dirac constant and defined as 10^{16} elements in a cm^3 and so forth.

An elucidation of the question concerning the limitation of a function to a large degree is based upon an analysis of the physical and logical essence of the studied process, its links and the dependences upon the absolute and relative limits in the investigated area of knowledge. A formal analysis of the differential growth function $\phi(t)$ can serve as an indication of the limit's existence. In the case of an asymptotic approximation to the limit, the function $\phi(t)$ will obviously move toward zero.

The secondary questions with an affirmative answer on the existence of a development limit include: what is the amount of this limit and what is the nature of the approach to the limit value of the studied function.

The solution of the first question often develops into independent forecast research. For example, a forecast of the limit computer speed with the set circuitry, the forecast for maximum power with the given fuel and overall dimensions of an engine, and so forth. The question is more simply solved in the case that the limit is absolute, relative or calculated.

The nature of the approach to the limit is also basically determined by logical analysis of the occurring process. A predominant number of technical and physical indicators for the processes move asymptotically toward their limits. At the same time certain volume indicators in economic forecasting can change over to limit values at a point. For example, the growth of the production volume for a certain product with limited capacity of raw material supplies or a sharp decline in the increase of sales with the complete saturating of the market. Such processes are significantly more difficult to extrapolate than are the asymptotic ones since the nature of their development, as a rule, is determined by external factors and is not brought out in analyzing the retrospective data about the process.

The soundness of the research and the degree of considering not only the short-term but also the long-term trends in the research are of great significance in correctly determining the constraints of the indicator and the nature of the move toward it.

R. Ayres [46] has pointed out that by using disaggregative analysis it is usually impossible to predict the appearance of even comparatively simple innovations (if they alter the configuration of the system), since the limit values for the efficiency of any class of systems (instruments) can be determined by extrapolation only on the level of individual components. Major inventions usually change the configuration of the components to such a degree that this in no way can be predicted ahead of time.

An aggregative approach is also aimed at an analysis and forecasting of a wide class of systems. In an extrapolation the aggregative approach has been embodied in the envelope method which will be taken up below. In applying this method the limits assume a more generalized sense and can be set more correctly and thereby more accurately reflect the future development of the process.

Hence, in setting development limits for any forecasted process it is always essential to correctly choose the level of analysis aggregating.

The third question of the listed basic questions about the nature of the process relates to the existence of bending points. This is also very essential in elaborating the various forecasts. In a number of studies the bending point and its position on the time scale has been chosen as the criterion for the perspective development of one or another scientific or technical area. Thus, in [41] for each area from the possible solutions of a technical (technological) problem a curve is constructed for patenting dynamics. Then, the presence of a bending point is shown on the curve corresponding to each area. In the event that patenting dynamics by the present has already passed its bending point, the given area for the development of technology (production methods) is considered relatively unpromising.

In formal terms the bending point can be established from the zero value of the second derivative of an approximating function with the subsequent change in the sign of this derivative. On the graph of $\phi(t)$, the bending point is shown in the form of an extremum so that the determining of it on the retrospective segment formally does not provide any problem. In the case of significant random fluctuations superimposed on the process, it is advisable to carry out its preliminary smoothing by one of the methods described in 2.4.3.

It is significantly harder to predict the appearance and location of a bending point in the future. Analysis of the essence of a developing process in this instance should be carried out in the direction of elucidating the factors which substantially influence the growth rate of the process and then an analysis of the trends of their change in the future.

An analysis of the bending points is substantially facilitated with the availability of data on limits in the development of the process and the nature of the move toward zero. In this instance the presence of a bending point can be considered set and its location is determined after choosing the parameters in the functional description.

As for the fourth question it must be pointed out that among the functions used for extrapolation in forecasting, only very few possess the property of symmetry. Basically these are logistical curves with a central symmetry relative to the bending point. Linear development laws possess a symmetry relative to any point but for them this property is not of particular significance.

The formal setting of symmetry can be carried out by an iterative comparison of the final differences of the function to the left and right of certain points which gradually approach a possible point corresponding most closely to the symmetry center. As the symmetry criterion relative to a certain point k (the symmetry center), it is possible to propose the following expression for the amount of the standard deviation of the final differences of an empirical series with a constant spacing to the right and left of the point:

$$S_k = \frac{2}{n}\sum_{i=1}^{n/2}(\Delta y_{k+i} \pm \Delta y_{k-i})^2, \tag{2.35}$$

where n--the total number of points in the segment of the curve studied for symmetry.

A plus sign is taken for axial symmetry and a minus sign for central symmetry. Then the presence of symmetry can be determined by the expression

$$\min_k \{S_k\} \leqslant \varepsilon, \tag{2.36}$$

where ε--the given positive number determining the limit for asymmetry.

The last question on the nature of the process concerns the elucidation of development constraints for the process but not in terms of amount but for time. This involves the elucidation and consideration in extrapolation of moments for the occurrence of certain events which bring about an end to the process or its transition to a different quality. As an example one might give the necessity of ending the production process for a certain product by a certain date as set by a higher-level plan or forecast.

In and of itself the setting of a time restriction on the process and a limitation point on a time scale can be the result of a forecast or a plan. Nevertheless the consideration of this factor is essential in extrapolating the development trends of parameters which describe this process or are linked to it.

Theoretically it is impossible to speak about a process (with the exception of the scale of the universe) as unlimited in time, and for this reason here it is a question of a relative limitation on the forecast's lead time. If the time of development (existence) for a process greatly exceeds the forecast lead time, then it is possible to speak about a process which does not have constraints in terms of development time. Otherwise it is essential to determine this constraint and consider it in choosing the extrapolating function.

Having analyzed the basic questions arising in the stage of choosing the type of function for an extrapolation of the studied process and the possible approaches to resolving them, let us move on to an examination of those functions which preferably should be used in a forecast extrapolation and certain demands which must be made on them. In [46] the demands are given which are made on the approximating curve: morphological simplicity, smoothness, symmetry and mathematical simplicity. As a whole it is possible to agree with such demands having put symmetry in last place and mathematical simplicity, having examined this in greater detail, in first place.

What do we understand by mathematical simplicity? This is the minimum possible number of terms in a formula; the minimum possible degree of an independent variable; a linear ascent of the coefficients; continuity; a minimal number of extremums and bending points. Such requirements are met by the standard functions which are most widespread in extrapolation:

a) Linear $y = a + bt$ (with $b = 0$, $y = a$--stable state);

b) Parabola $y = a + bt + ct^2$ (with $c > 0$--a growth function, with $c < 0$--an extremal function);

c) Step function $y = at^b$;

d) Exponential function $y = ae^{bt}$;

e) Modified exponent $y = k - ae^{-bt}$;

f) Logistical (S-shaped) curve $y = \frac{k}{1 + be^{-ct}}$;

g) Hyperbolic function $y = a + \frac{b}{c+t}$.

The various sources give different lists of recommended curves. The lists given here, in our view, represents that essential minimum which covers a predominant portion of the needs arising in extrapolating trends in economic and scientific-technical forecasting.

This list does not include functions making it possible to approximate periodic processes. The problem is that harmonic analysis of periodic processes represents a rather independent problem. Moreover, in forecasting problems in the area of the BTS, the necessity of extrapolating periodic processes virtually does not arise (with rare exceptions).

In a general form, the sequence of steps in selecting the type of extrapolating function for the set empirical series can be as follows: Smoothing the empirical series, attempts at linear fitting; in the case of failure the constructing of differential growth functions; in the case of success immediately a logical analysis of the essence of the process using the scheme given in the given section and the choice of the type of function for the extrapolation.

2.4.5. Calculation of Parameters for the Extrapolating Function

Thus, as a result of solving the previous stages in the forecast research a definite type of function has been found for extrapolating an empirical series and its analytical presentation has been given containing a series (or one) of unknown parameters. In the following stage it is essential, in using the empirical series, to choose (calculate) the values of these parameters which ensure optimum approximation in a certain sense.

As the optimality criterion ordinarily one uses one or another measure of the deviations of the points in the empirical series from the approximating function. Each of the possible criteria for approximation optimality has its corresponding method for determining the curve parameters. Let us examine the basic methods having paid most attention to the least square method as the most widely found.

The averages method is based upon the minimization of the algebraic total of point deviations from an approximating curve. In this case the optimality criterion is written in the following form:

$$S = \sum_{i=1}^{n} y_i - f(x_i, a_1, a_2, \ldots, a_m) \rightarrow \min, \qquad (2.37)$$

where y_i, x_i--the ordinate and abscissa of point i of the series; $a_1, a_2, \ldots, a_m$--the parameters of the approximating curve.

In practice this method is realized in the following manner. All the points of the empirical series (n > m) are divided evenly into m groups and for each of them the total of the deviations of the type (2.37) is reduced to zero. As a result one obtains a system of m equations with m unknown parameters $a_1, a_2, \ldots, a_m$ which are determined by solving this system. With a linear incorporation of the parameters into the curve's formula, the system is linear and is solved by one of the known methods.

It is essential to point out that the results of the method of averages depend substantially upon the method for grouping the points. Practice shows that the most rational is a grouping method whereby the groups are made up of points which increase successively with the independent variable.

With a linear function of f(x) = ax+b, we obtain a following type of system (for even n):

$$\begin{cases} a\sum_{i=1}^{n/2} x_i + b\frac{n}{2} = \sum_{i=1}^{n/2} y_i; \\ a\sum_{i=n/2+1}^{n} x_i + b\frac{n}{2} \sum_{i=n/2+1}^{n} y_i. \end{cases} \tag{2.38}$$

In adding together both equations of the system and having divided them by n, we obtain

$$a\bar{X} + b = \bar{Y}. \tag{2.39}$$

From (2.38) and (2.39) we determine the parameters a and b.

$$a = \frac{\frac{2}{n}\sum_{i=1}^{n/2} y_i - }{\frac{2}{n}\sum_{i=1}^{n/2} x_i - \bar{X}} \quad \text{and } b = \bar{Y} - a\bar{X}. \tag{2.40}$$

Analogously it is possible to determine the coefficients for the function of three parameters and generally for m parameters.

The method of averages and the corresponding criterion provide satisfactory results for forecasting for sufficiently smooth (or smoothed) numerical series and significantly poorer ones with substantial random components.

The source [43] gives another series of methods based on criteria which differ from (2.37):

$$S_1 = \frac{\sum_{i=1}^{n} |y_i - f(x_i, a_1, \ldots, a_m)|}{n-2} \to \min; \tag{2.41}$$

$$S_2 = \sqrt{\frac{\sum_{i=1}^{n} [(y_i - f(x_i, a_1, a_2, \ldots, a_m)]^2}{n-2}} \to \min; \tag{2.42}$$

$$\chi = \sqrt{\frac{\sum_{i=1}^{n} \left[\frac{y_i - f(x_i, a_1, \ldots, a_m)}{f(x_i, a_1, \ldots, a_m)}\right]^2}{n-1}} \to \min. \tag{2.43}$$

We will not stop to analyze in detail the advantages and drawbacks of each of these criteria but we would merely point out that the results obtained from each of them vary and in certain instances are even contradictory.

Let us examine one other extrapolation method which contains elements of graphic constructs and for this reason is simplest and most approximate. In the given instance the formal optimality criterion is absent and is replaced by a certain qualitative concept of visual proximity.

The points of the empirical series are plotted on the coordinate plane XOU, after which a smooth curve is drawn on it in passing most closely to all the points. From the type of curve its order m is estimated and according to this the point (m + 1) is chosen on it. These points are distributed approximately evenly along the entire interval of the set of the series. The abscissas and ordinates of these points make it possible to compile a system from the (m + 1) equation for determining the values of the (m + 1) curve parameter analogously as is done for the method of averages.

It must be pointed out that this informal method is good for preliminary estimate calculations and simple forms of curves. The use of human intuition and the ability to draw smooth averages often gives it very satisfactory results.

The least square method is the most widespread method in forecast extrapolation. Its merits are the relative simplicity of realization (for a number of functions it has been brought to an analytical presentation of coefficients). The method smooths out well the random noise in describing a trend and it makes it possible to obtain unbiased and consistent estimates for all the parameters a_0, a_1, ..., a_m. In the most widespread case of the linear incorporation of the parameters into the trend formula, the parameter estimates using the least square method are also effective.

The formulating of the least square method comes down to the following. If all the measurements for the values of the function y_1, y_2, ..., y_n are made with equal precision, then the estimates of the function parameters are determined by the condition

$$S = \sum_{\kappa=1}^{n} [y_\kappa - f(x_\kappa;\ a_0,\ a_1,\ \ldots,\ a_m)]^2 \rightarrow \min. \qquad (2.44)$$

If the measurements have been made with different accuracy (with different variance), then weights are introduced which are inversely proportional to the assumed variance ratios:

$$S = \sum_{\kappa=1}^{n} [y_\kappa - f(x_\kappa;\ a_0,\ a_1,\ \ldots,\ a_m)]^2\, w_\kappa \rightarrow \min; \qquad (2.45)$$

$$w_1 : w_2 : \ldots : w_\kappa : \ldots : w_n = \frac{1}{\sigma_1^2} : \frac{1}{\sigma_2^2} : \ldots : \frac{1}{\sigma_\kappa^2} : \ldots : \frac{1}{\sigma_n^2}.$$

If for each value of the independent variable x_k there are several measurements m_k for the function of values and the average arithmetic value of the results is taken as y_k, then the numbers of the measurements in the series $w_k = m_k$ can serve as the measurement weights.

In the general case the least square method can be extended to the case of the function of M variables:

$$S=\sum_{k=1}^{n}[y_k-f(x_{1k}, x_{2k}, \ldots, x_{Mk}, a_0, a_1, \ldots, a_m)]^2 w_k \to \min.$$

The most widespread form of using the method in economic research is (2.44). In forecast research it is possible to use the form (2.45) when one considers the so-called discounting of data (a reduction in value and accuracy) as one moves into the past using the corresponding discounting coefficients w_k. In a general form the realization of the formulas of the type (2.44) or (2.45) leads to the solving of a system of equations of the sort:

$$\frac{\partial S}{\partial a_0}=0;\ \frac{\partial S}{\partial a_1}=0;\ \ldots;\ \frac{\partial S}{\partial a_n}=0. \qquad (2.46)$$

In the case that the parameters of a_k are introduced linearly into the accepted formula $f(x, a_0, a_1, \ldots, a_m)$, then the system of equations of (2.46) will also be linear. Let us examine several specific types of functions.

The linear function: $y = ax+b$. In this instance the system of equations (2.46) assumes the form:

$$\begin{gathered} a\sum_{1}^{n} x_k + bn = \sum_{1}^{n} y_k; \\ a\sum_{1}^{n} x_k^2 + b\sum_{1}^{n} x_k = \sum_{1}^{n} x_k y_k. \end{gathered} \qquad (2.47)$$

However, since with this method a straight line always passes through the average values of the $\bar{x}$ and $\bar{y}$ coordinates, it is advisable to construct it as a single-parameter one of the type

$$y-\bar{y}=a(x-\bar{x});$$

$$\bar{x}=\frac{\sum_{}^{n} x_k w_k}{\sum_{}^{n} w_k};\quad \bar{y}=\frac{\sum_{}^{n} y_k w_k}{\sum_{}^{n} w_k};$$

$$a=\frac{\overline{xy}-\bar{x}\bar{y}}{\overline{x^2}-(\bar{x})^2};$$

$$\overline{x^2}=\frac{\sum x_k^2 w_k}{\sum w_k} \quad \text{and} \quad \overline{xy}=\frac{\sum x_k y_k w_k}{\sum w_k}.$$

The programming of these formulas and computer calculation does not represent any difficulty.

The quadratic function $y = ax^2+bx+c$ leads to a system of equations of the type

$$\left\{\begin{aligned} aS_4 + bS_3 + cS_2 &= \sum_{\kappa=1}^{n} y_\kappa x_\kappa^2 w_\kappa; \\ aS_3 + bS_2 + cS_1 &= \sum_{\kappa=1}^{n} y_\kappa x_\kappa w_\kappa; \\ aS_2 + bS_1 + cS_0 &= \sum_{\kappa=1}^{n} y_\kappa w_\kappa; \end{aligned}\right. \tag{2.48}$$

$$S_p = \sum_{\kappa=1}^{n} x_\kappa^p w_\kappa \ (p = 0, 1, 2, 3, 4).$$

The solving of the given system of equations also is rather easy to carry out on a computer, however it can be further simplified by shifting the start of the calculation to the point $\bar{x}$ (for the case $w_k = 1$) and measuring the function of y at points which are symmetrical to $\bar{x}$ on the left and right. Then,

$$\begin{gathered} y = a\,(x - \bar{x})^2 + b\,(x - \bar{x}) + c; \\ a = \frac{1}{D}\left[\sum_{\kappa=1}^{n} y_\kappa\,(x_\kappa - \bar{x})^2 - S_2\bar{y}\right]; \\ b = \frac{1}{S_2}\sum_{\kappa=1}^{n} y_\kappa\,(x_\kappa - \bar{x}); \\ c = \bar{y} - \frac{S_2 a}{n};\ D = S_4 - \frac{S_2^2}{n};\ S_p = \sum_{\kappa=1}^{n} (x_\kappa - \bar{x})^p\ (p = 0,2,4). \end{gathered} \tag{2.49}$$

Parabolas are frequently employed in forecast extrapolation and the calculating of them using the formulas of (2.49) has been substantially simplified.

At the conclusion of this section we will become familiar with still another method for determining the parameters of an extrapolation function and this is based on the uneven consideration of the importance of retrospective information. In a majority of practical cases, of great significance for the forecasting of a process is the information which directly describes the process at moments of time standing closer to the future (that is, present moments) and the further we move into the past history of the process the less valuable the information will be for the future. In the least square method, this devaluing of information can be considered by discounting coefficients w_k (2.45).

Let us examine the method which has been named exponential smoothing because of the exponential nature of discounting the retrospective values of variables. Its essence is that the time series is smoothed using a weighted moving average and here the weights are subordinate to an exponential law. The weighted moving average with the exponentially distributed weights characterizes the significance of the process at the end of the smoothing interval, that is, is the mean characteristics for the last terms of the series. Such a weighted moving average can be a forecasting tool.

The exponential smoothing method is employed for forecasting series which are described by polynomials:

$$y_t = b_0 + b_1 t + \frac{b_2}{2!} t^2 + \cdots + \frac{b_p}{p!} t^p + \varepsilon_t = \sum_{j=0}^{p} \frac{b_j}{j!} t^j + \varepsilon_t. \qquad (2.50)$$

A forecast for moments of time $t = (T + \ell)$, where $\ell = 1, \ldots, L$, is compiled by weighing the observations of the series y_t in such a manner that greater weights are assigned to the later observations in comparison with the earlier ones. The forecast of the dynamic series y_t for the time period $T + \ell$ can be constructed using an expansion into a Taylor series:

$$\overset{\bullet}{y}_{T+l} = y_t^0 + l y_t^1 + \frac{l^2}{2!} y_t^2 + \cdots + \frac{l^\kappa}{\kappa!} y_t^\kappa + \cdots + \frac{l^p}{p!} y_t^p,$$

where y_t^k--derivative k taken at the moment of time t.

According to the theorem proved by Brown and Mayer, any k derivative (k = 0, 1, ..., p) of an equation can be expressed by linear combinations of exponential means up to the (p + 1) order. The calculating of the recurrent corrections for the estimates of coefficients for the equation (2.50) here is the basic aim of exponential smoothing.

An exponential first order mean for the series y_t is the name given to:

$$S_t^{[1]}(y) = \alpha \sum_{i=0}^{n} (1 - \alpha)^i \, y_{t-i}, \qquad (2.51)$$

where α--the smoothing parameter $(0 < \alpha < 1)$.

The expression $S_t^{[\kappa]}(y) = \alpha \sum_{i=0}^{n} (1 - \alpha)^i S_{t-1}^{[\kappa-1]}(y)$ is called an exponential mean of the k order (k = the number of terms in the polynomial).

Brown derived the following recurrent formula for determining the exponential mean

$$S_t^{[\kappa]}(y) = \alpha S_t^{[\kappa-1]}(y) + (1 - \alpha) S_{t-1}^{[\kappa]}(y). \qquad (2.52)$$

The new exponential mean equals the previous one plus the portion of α from the difference between the new observations and the previous smoothed values of the series. The function (2.52) is a linear combination of all past observations. The weights assigned to the previous terms of the series diminish in a geometric progression.

From formula (2.52) it can be seen that for smoothing the series it is essential to set an initial amount of the exponential mean. For this it is possible to use the special formulas worked out by Brown. In particular, for a linear model the initial conditions are determined in the following manner:

$$S_0^{[1]}(y) = b_0 - \frac{1 - \alpha}{\alpha} b_1;$$

$$S_0^{[2]}(y) = b_0 - \frac{2(1 - \alpha)}{\alpha} b_1. \qquad (2.53)$$

The coefficients b_0 and b_1 in (2.53) are the coefficients for a linear equation of a time trend obtained by the least square method.

For a quadratic model the initial conditions are set from the expressions:

$$\begin{aligned} S_0^{[1]}(y) &= b_0 - \frac{1-\alpha}{\alpha} b_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2} b_2; \\ S_0^{[2]}(y) &= b_0 - \frac{2(1-\alpha)}{\alpha} b_1 + \frac{(1-\alpha)(3-2\alpha)}{\alpha^2} b_2; \qquad (2.54) \\ S_0^{[3]}(y) &= b_0 - \frac{3(1-\alpha)}{\alpha} b_1 + \frac{3(1-\alpha)(4-3\alpha)}{2\alpha^2} b_2. \end{aligned}$$

The coefficients b_0, b_1 and b_2 in (2.54) are the coefficients of a quadratic equation for a time trend obtained by the least square method.

In order to express the coefficients of a linear equation which includes only the first two terms of the model of (2.50), on the basis of the Brown-Mayer theorem it is possible to obtain a system of equations connecting the estimates for the coefficients b_0 and b_1 with the exponential means $S_t^{[1]}(y)$ и $S_t^{[2]}(y)$:

$$\begin{aligned} S_t^{[1]}(y) &= \hat{b}_0 + \frac{1-\alpha}{\alpha} \hat{b}_1; \\ S_t^{[2]}(y) &= \hat{b}_0 + \frac{2(1-\alpha)}{\alpha} \hat{b}_1. \qquad (2.55) \end{aligned}$$

In solving the system for $\hat{b}_0$ and $\hat{b}_1$, we obtain

$$\begin{aligned} \hat{b}_0 &= 2S_t^{[1]}(y) - S_t^{[2]}(y); \\ \hat{b}_1 &= \frac{\alpha}{1-\alpha} [S_t^{[1]}(y) - S_t^{[2]}(y)]. \qquad (2.56) \end{aligned}$$

For a quadratic model we obtain the following system of three equations with three unknowns:

$$\begin{aligned} S_t^{[1]}(y) &= \hat{b}_0 - \frac{1-\alpha}{\alpha} \hat{b}_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2} \hat{b}_2; \\ S_t^{[2]}(y) &= \hat{b}_0 - \frac{2(1-\alpha)}{\alpha} \hat{b}_1 + \frac{(1-\alpha)(3-2\alpha)}{\alpha^2} \hat{b}_2; \qquad (2.57) \\ S_t^{[3]}(y) &= \hat{b}_0 - \frac{3(1-\alpha)}{\alpha} \hat{b}_1 + \frac{3(1-\alpha)(4-3\alpha)}{2\alpha^2} \hat{b}_2. \end{aligned}$$

Hence

$$\begin{aligned} \hat{b}_0 &= 3[S_t^{[1]}(y) - S_t^{[2]}(y)] + S_t^{[3]}(y); \\ \hat{b}_1 &= \frac{\alpha}{2(1-\alpha)^2} [(6-5\alpha) S_t^{[1]}(y) - 2(5-4\alpha) S_t^{[2]}(y) + \\ &\quad + (4-3\alpha) S_t^{[3]}(y); \qquad (2.58) \\ \hat{b}_2 &= \frac{\alpha^2}{(1-\alpha)^2} [S_t^{[1]}(y) - 2S_t^{[2]}(y) + S_t^{[3]}(y)]. \end{aligned}$$

The forecast for a linear model is made using the formula

$$y^*_{t+\ell} = \hat{b}_0 + \hat{b}_1 t. \qquad (2.59)$$

The forecast error is calculated from the expression

$$S_{y^{\bullet}_{t+l}} = S_{\tilde{y}(t)} \sqrt{\frac{\alpha}{(2-\alpha)^3}[1+4(1-\alpha)+5(1-\alpha)^2+ \cdots\!\rightarrow \rightarrow\!\cdots +2\alpha(4-3\alpha)l+2\alpha^2 l^2]}. \qquad (2.60)$$

A forecast for a quadratic model is calculated from the formula

$$y^{\bullet}_{t+l} = \hat{b}_0 + \hat{b}_1 l + \frac{1}{2}\hat{b}_2 l^2 \qquad (2.61)$$

with an error

$$S_{y^{\bullet}_{t+l}} \approx S_{\tilde{y}(t)} \sqrt{2\alpha + 3\alpha^2 + 3\alpha^3 l}. \qquad (2.62)$$

In an analogous manner it is possible to obtain estimates for the remaining coefficients of the polynomial (2.50).

In constructing forecasts using the exponential smoothing method, one of the basic problems is the choice of the value of the smoothing parameter α. Naturally, with the different values of α the forecast results will vary. The weight of the observation left for k periods from the observed moment equals $\alpha(1-\alpha)^k$. If it is certain that the initial conditions are reliable, then one must use a small amount of the smoothing parameter ($\alpha \simeq 0$). If there is not sufficient certainty of the reliability of the initial conditions, then one must use a large amount and this leads to considering basically the influence of late observations in the forecast.

2.4.6. Functions With a Flexible Structure in Trend Extrapolation

An examination of the various methods for choosing the type of approximating function and the procedures accompanying them indicates that this process is essentially based on the intuition and experience of the forecaster and contains a mass of informal qualitative estimates and criteria. Naturally because of this there has been a desire on the part of the researchers to formalize this process, to describe it with clear quantitative criteria and to elaborate the mathematical procedures for its optimization.

As an example of such studies one might give the ideas developed in a number of workd [19, 20] devoted to the theory and practice of using so-called flexible structure functions (FGS).

In contrast to ordinary functions described analytically, it is proposed that use be made of functions the appearance of which can be altered depending upon the values incorporated in describing the parameter function. In using functions with a flexible structure in forecasting in the process of minimizing the errors between the FGS values and the points of the empirical series, not only should the optimum values of the coefficients for the approximating function be chosen but also its optimal type (in accord with the accepted criterion). This is the basic idea of the method. Below is given its formal description.

The initial process (a single variable function) is presented in the form

$$z(x) = F(x) + R(x), \qquad (2.63)$$

where F(x)--the approximate model of the process (a description using the FGS);
R(x)--the remainder (a certain function of the approximation accuracy).

In the most general form the FGS for a single independent variable can be written:

$$F(x) = A_0 + \sum_{j=1}^{n} A_j \frac{\delta_j(x - x_0)}{D}, \qquad (2.64)$$

where n--the fixed natural number;
x_0--the initial value of the independent variable in the designated interval;
$A_0, A_1, \ldots, A_n$--the constant active parameters;
D--the special determinant of order n;

$$D = \begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ r_1 & r_2 & r_3 & \ldots & r_n \\ r_1^2 & r_2^2 & r_3^2 & \ldots & r_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ r_1^{n-1} & r_2^{n-1} & r_3^{n-1} & \ldots & r_n^{n-1} \end{vmatrix}, \qquad (2.65)$$

in which $r_1, r_2, \ldots, r_n$ are real or complex paired numbers;
$\delta_j(x - x_0)$--functions obtained from the determinant D by substituting the elements of line j for exponents of the type $1/r_\nu[e^{r_\nu(x-x_0)}-1]$, $\nu = 1, 2, \ldots, n$.

Thus, the FGS depends upon the independent variable x and the values of (2n + 2) parameters (x_0; n; $A_0, \ldots, A_n$; $r_1, \ldots, r_n$) and in such a manner that the change in them can cause a change in the very structure of F(x). The structure F(x) is basically influenced by the values r_j (j = 1, 2, ..., n). In the event that all the parameters of $r_1, r_2, \ldots, r_n$ are real, different and not equal to zero then the function F(x) is determined, real and continuous on any segment of the independent variable. Actually in the given instance it will be a linear combination of exponents which grow without limitation with $r > 0$ and diminish without limitation for $r < 0$.

If n--an even number (n = 2p) and the parameters $r_1, r_2, \dots, r_n$ are complex, paired numbers of the type $r_\mu = \alpha_\mu + \beta_\mu i$ and $r_{\mu+1} = \alpha_\mu - \beta_\mu i$, where $i = \sqrt{-1}$, $\beta_\mu \neq 0$, $\mu = 1, 2, \dots, p$ and α_μ are real and different, then the function F(x) according to the formula (2.63) is determined, real and continuous on any segment of the independent variable and can be represented in the following form:

$$F(x) = \Phi(x) = B_0 + \sum_{\mu=1}^{p} e^{\alpha_\mu (x-x_0)} [B_\mu \cos \beta_\mu (x-x_0) + C_\mu \sin \beta_\mu (x-x_0)],$$

where B_0, B_μ, C_μ--constant real parameters depending upon $A_0, A_1, \dots, A_n$; $\alpha_1, \alpha_2, \dots, \alpha_p$; $\beta_1, \beta_2, \dots, \beta_p$.

As is seen, in this case F(x) will contain terms of a harmonic nature which are amplitude modulated by the exponential functions.

In the event that all the terms $r_1, r_2, \dots, r_n$ move toward zero with a transition to the limit, the function F(x) can be reduced to the type

$$F(x) = T_n(x) = A_0 + A_1 (x-x_0) + \dots + A_n \frac{(x-x_0)^n}{n!},$$

that is, we obtain a step polynomial.

From an examination of the different variations for setting the parameters $r_1, r_2, \dots, r_n$ it becomes apparent that by changing them it is possible to alter the structure of F(x) in rather broad limits, representing it in the form of a step polynomial, a trigonometric polynomial, or an aggregate of exponents, or, finally, any combination of the listed structures. This is a very enticing property of the FGS from the viewpoint of using them for trend extrapolation.

It remains to solve the question about the method to seek out the optimum values of $r_1, r_2, \dots, r_n$. Let us examine a case when the function z(x) is differentiatable (n + 1) times in a segment containing the initial point $x = x_0$. Then formula (2.64) can be reduced to the type

$$F(x) = z(x_0) + \sum_{j=1}^{n} z^j(x_0) \frac{\delta_j (x-x_0)}{j!}, \tag{2.66}$$

where $z^j(x_0)$--the derivative of order j for the initial function at the starting point;

D--the determinant described by (2.65) where $r_1, r_2, \dots, r_n$ are viewed as roots of a certain base equation:

$$\begin{gathered} r^n + a_{n-1} r^{n-1} + a_{n-2} r^{n-2} + \dots + a_1 r + a_n = 0; \\ a_0 = (-1)^n r_1 r_2 \dots r_n; \\ a_1 = (-1)^{n-1} (r_1 r_2 \dots r_{n-1} + r_2 r_3 \dots r_n); \\ \dots\dots\dots\dots\dots\dots \\ a_{n-1} = (-1)^{n-(n-1)} (r_1 + r_2 + \dots + r_n). \end{gathered} \tag{2.67}$$

The functions $\delta_j(x-x_0)$ are determined analogously to (2.64).

An analytical expression of the remainder in formula (2.63) has the following form:

$$R(x) = \int_{x_0}^{x} \int_{x_0}^{t} \eta(\tau) \frac{\Delta_n(t-\tau)}{D} d\tau\, dt. \tag{2.68}$$

The term of this formula $\Delta_n(t-\tau)$ can be obtained from (2.65) in substituting the last n line of the determinant D for the line of the exponents of the type $e^{r_\nu(t-\tau)}$, ($\nu = 1, 2, \ldots, n$).

The other term of (2.68) can be defined from the equation

(2.69)

The idea of an optimal approximation in using the FGS comes down to minimizing the remainder R(x) and to establishing such values of the parameters $a_0, a_1, \ldots, a_{n-1}$ that the value of the remainder in each point of the segment does not exceed a certain set amount (the approximation error) and this also determines the type of F(x) which is the solution to the problem.

In [20] as an example they examine a case of even approximation

$$\left| z(x) - A_0 - \sum_{j=1}^{n} A_j \frac{\delta_j(x-x_0)}{D} \right| < \gamma_0, \tag{2.70}$$

where γ_0 = const. > 0, $\alpha < x_0 \leqslant x < x_k \leqslant b$.

If the parameters of A_j are taken according to (2.66), then we obtain an expression of the approximation error in the form

$$\varepsilon(x) = z(x) - \left[z(x_0) + \sum_{j=1}^{n} z^{(j)}(x_0) \frac{\delta_j(x-x_0)}{D} \right]. \tag{2.71}$$

The unknowns in (2.71) are the parameters $r_1, r_2, \ldots, r_n$ which can be determined in the process of minimizing $\varepsilon(x)$. With an even approximation the errors values of $\varepsilon(x)$ at the extremal points $M_1, M_2, \ldots, M_n$, at the initial point M_0 and the end point $M_{n+1} = M_k$ are the same for the modulus and alternating in sign:

$$\begin{gathered} \varepsilon'(x_0) = \varepsilon'(x_1) = \cdots = \varepsilon'(x_n) = 0; \quad x_0 \leqslant x < x_\kappa; \\ \varepsilon(x_p) = -\varepsilon(x_{p+1}),\ p = 1, 2, \ldots, n;\ |\varepsilon(x_j)| = \\ = N,\ j = 0, 1, 2, \ldots, n+1;\ |\varepsilon(x)| < N. \end{gathered} \tag{2.72}$$

Depending upon the selected system of base functions and the different values of the parameters $a_0, a_1, \ldots, a_{n-1}$ in the base equation (2.67) in the process of minimization of (2.71) various structures of F(x) can be obtained, for example, a Taylor step polynomial, a Chebyshev polynomial or a trigonometric series.

The other method of even approximation can be termed integral. It is based upon the use of the formula for the remainder in (2.68):

$$\varepsilon(x) = z(x) - F(x) = R(x) = \int_{x_0}^{x}\int_{x_0}^{t} \eta(\tau)\frac{\Delta_n(t-\tau)}{D}\,d\tau\,dt, \tag{2.73}$$

Then the first part of the conditions in (2.72) can be expanded on this basis. In particular, the conditions of the extremums $\varepsilon(x)$ at points M_1, M_2, ..., M_n are written: $\varepsilon'(x_1) = \varepsilon'(x_2) = \ldots = \varepsilon'(x_n) = 0$, or considering (2.73)

$$\int_{x_0}^{x_1} \eta(\tau)\frac{\Delta_n(x_1-\tau)}{D}\,d\tau = \cdots = \int_{x_0}^{x_n} \eta(\tau)\frac{\Delta_n(x_n-\tau)}{D}\,d\tau = 0. \tag{2.74}$$

The second part of the conditions in (2.72) will assume the form

$$\int_{x_0}^{x_p}\int_{x_0}^{t} \eta(\tau)\frac{\Delta_n(t-\tau)}{D}\,d\tau\,dt = -\int_{x_0}^{x_{p+1}}\int_{x_0}^{t} \eta(\tau)\frac{\Delta_n(t-\tau)}{D}\,d\tau\,dt. \tag{2.75}$$

If for a certain function of z(x) it is possible to choose a value of n and the parameters of a_0, ..., a_{n-1} in order to fulfill the condition

$$\eta(x) = z^{(n+1)}(x) + a_{n-1}z^{n}(x) + \ldots + a_1 z^{(2)}(s) + a_0 z^{(1)}(x) \equiv 0,$$

then F(x) will provide an even approximation with a zero error factor. Unfortunately, the nonlineality of incorporating the coefficients a_0, ..., a_{n-1} in the system of (2.74), (2.75) substantially impedes its practical utilization.

In the instance of a small segment $[x_0, x_k]$, if n is sufficiently large, it is possible to assume

$$e^{r_\nu(t-\tau)} = + r_\nu(t-\tau) + \frac{r_\nu^2(t-\tau)}{2!} + \cdots + \frac{r_\nu^{n-1}(t-\tau)}{(n-1)!}; \tag{2.76}$$
$$\nu = 1, 2, \ldots, n, \; x_0 < \tau < t < x < x_k,$$

then the equality is valid

$$\frac{\Delta_n(t-\tau)}{D} = \frac{(t-\tau)}{(n-1)!}$$

and the system of equations in (2.74) and (2.75) is linear relative to a_0, a_1, ..., a_{n-1}:

$$\int_{x_0}^{x_1} \eta(\tau)(x_1-\tau)^{n-1}\,d\tau = \cdots = \int_{x_0}^{x_n} \eta(\tau)(x_n-\tau)^{n-1}\,d\tau = 0; \tag{2.77}$$

$$\int_{x_0}^{x_p}\int_{x_0}^{t} \eta(\tau)(t-\tau)^{n-1}\,d\tau\,dt = -\int_{x_0}^{x_{p+1}}\int_{x_0}^{t} \eta(\tau)(t-\tau)^{n-1}\,d\tau\,dt. \tag{2.78}$$

In practice it is not known ahead of time whether or not (2.76) is fulfillable, and for this reason in the instance that it is possible to formally solve the system (2.77) and (2.78) relative to n of the parameters a_0, a_1, ..., a_{n-1} and n of the coordinates x_1, x_2, ..., x_n, then the roots of r_1, r_2, ..., r_n are calculated and the correctness of (2.76) is checked.

In the general case the seeking out of the FGS parameters leads to a system of transcendental equations which, as a rule, do not have analytical solutions. Nevertheless a number of particular cases of the functions can be successfully approximated on the basis of the FGS.

In forecasting with the aid of the FGS, it must not be forgotten that this method is only one of the methods of formal extrapolation, that is, the optimum fitting of a series which does not contain elements of forecast extrapolation in the understanding that is presented in 2.4.1. The very idea of a method for automatically obtaining the type and parameters of an approximating function is extremely interesting and enticing, however from the viewpoint of practical implementation there still are numerous difficulties.

2.4.7. Envelope Extrapolation

Envelopes in forecasting reflect an aggregative approach in the analysis and forecasts of complex processes. As an examination of the developmental history of scientific and technical progress shows, this process can be broken down into two conditional components: evolutionary and revolutionary. The former is characterized by a relatively gentle change determined by the quantitative accumulation of knowledge in the process of researching, improving technology, production methods, materials quality and so forth. The second is determined by qualitative shifts or bursts, as certain foreign authors term them.

In elaborating the long-range forecasts of a scientific-technical or economic nature, the degree of aggregation should rise along with the increased lead time. It is possible to give a number of examples confirming this rule. R. Ayres [46] gives graphs for the increase in the effective energy of particle accelerators and the trends for the increased power of aircraft engines. From examining them it becomes apparent that a specific type of technical device operating on a certain principle causes the growth of the forecasted variable in a certain finite time interval and a certain finite interval on the scale of values. The dimensions of the time interval depend upon a large number of factors in the development of scientific and technical progress and these determine the time interval for the existence of the given type from the moment prior to which its appearance had not been prepared for by the entire history of development and up to the moment of the aging of the given type and the exhaustion of its possibilities. The interval on the scale of values is determined by certain relative limits caused by the essence and principles of the examined type.

An envelope provides an opportunity to disclose a general development trend in the forecasted variable and to assess the possible limits of its development and the nature of the approach to these limits. Here definite opportunities are opened up to predict not only the nature of evolutionary growth but also the probable technological or scientific bursts, shifts in development and the appearance of fundamental inventions which is the most important and most difficult task in forecasting scientific and technical progress.

As in the case in the problems of extrapolating numerical series, in envelope extrapolation it is possible to isolate the purely formal part and a number of aspects which make this extrapolation a forecast one. Let us examine the formal mathematical set up of the problem and the possible ways of solving it.

Let us accept as given the set P of experimental points and this set can be broken down into a certain number of subsets P_I, P_2, P_3, ..., P_n according to a certain parameter C. As C we will accept a certain generalized concept of the parameter which does not come down just to a numerical amount but can also be a certain nonquantitative characteristic such as a "type of device," "operating principle," "physical nature," "elemental base," and so forth. In the general instance the C parameter can be complex and represent the vector $C = |c_1, c_2, \ldots, c_k|$. As an example we might turn to [46], where an entire range of accelerator characteristics such as the type of particle trajectory, the method of beam focusing and so forth is used as the parameter in extrapolating accelerator capacity.

Let the designated subsets of points form on plane XOY a certain family of curves which we will call the C parameter family (Fig. 2.11). In the event that the C parameter has the nature of a continuous quantity, the envelope could be defined as a curve which at each point has a common tangent to one of the family of curves. In the case of a discrete nature for the change in C it is not possible to give a strict definition of the envelope as it is impossible to draw it uniformly. By an envelope we will understand the smoothest of all the possible curves touching all (or a majority) of the curves in the family.

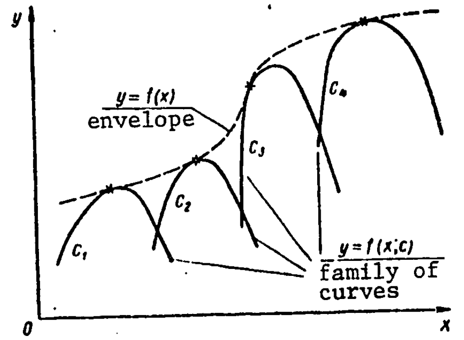


Fig. 2.11. Illustration of the concept of an envelope of a family of curves

Let us examine a possible method of determining it. Let each subset of points P_1, P_2, ..., P_m be approximated by the above-given procedures using dependences of the type:

$$\begin{aligned} y_1 &= f_1(a_{11}, a_{12}, \ldots, a_{1n}, x); \\ y_2 &= f_2(a_{21}, a_{22}, \ldots, a_{2n}, x); \\ y_m &= f_m(a_{m1}, a_{m2}, \ldots, a_{mn}, x), \end{aligned} \qquad (2.79)$$

where all the parameters of a_{ij} are determined in the approximation process and are considered known (for the sake of simplicity let us consider their number n as the same for all the functions and this does not in principle play an essential role). Let us assume that we are looking for the envelope equation in the form of k parametric curve of the type $Y = F(b_1, b_2, \ldots, b_k, x)$. For each point of its tangency with one of the curves of the family we will have two equations of the type

$$Y_e = y_e \quad \text{and} \quad \left(\frac{dY}{dx}\right)_e = \left(\frac{dy}{dx}\right)_e, \qquad (2.80)$$

where e = 1, 2, ..., m--the number of the tangency point.

In substituting the expressions for Y and y in these equations, we will obtain a system of 2m equations with the unknowns $x_1, x_2, \ldots, x_m$ and $b_1, b_2, \ldots, b_k$ of the type

$$\begin{gathered}
f_1(a_{11}, a_{12}, \ldots, a_{1n}, x_1) = F(b_1, b_2, \ldots, b_K, x_1);\\
\left(\frac{df_1(a_{11}, a_{12}, \ldots, a_{1n}, x)}{dx}\right)_{x=x_1} = \left(\frac{dF(b_1, b_2, \ldots, b_K, x)}{dx}\right)_{x=x_1};\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
f_m(a_{m1}, a_{m2}, \ldots, a_{mn}, x_m) = F(b_1, b_2, \ldots, b_K, x_m);\\
\left(\frac{df_m(a_{m1}, a_{m2}, \ldots, a_{mn}, x)}{dx}\right)_{x=x_m} = \left(\frac{F(b_1, b_2, \ldots, b_K, x)}{dx}\right)_{x=x_m}.
\end{gathered} \tag{2.81}$$

In the event that k = m, that is, the number of parameters in the envelope equation equals the number of curves in the family, we obtain a system of 2m equations with 2m unknowns, that is, in principle a solution is possible. If k > m, then there is no solution, but if k < m, then an approximate solution is possible.

However, it must be pointed out that since the variable x is included nonlinearly in the equations of the family, the system (2.81) is nonlinear in all practical instances. In this regard the solution to it is impossible in an analytical form in real situations and, in addition, the numerical solution methods also may not lead to positive results. On the other hand, such a formal solution is very critical toward mistakes in setting the points of the initial series or to the random components present in an empirical description of the process.

At present, there is no known positing of the problem of constructing an envelope for a family of curves in an approximate variation as has been done for the extrapolation of dynamic series, for example, by the least square method. Because of this the approximate graphic-analytical method can be given as the basic method for constructing envelopes. It can be realized in the following manner.

The approximating curves determined in one way or another for the subsets of points from the family $P_1, P_2, \ldots, P_m$ are dipicted on the on the coordinate plane XOY. After this a smooth curve is drawn close to the above-given concept of an envelope. In the given instance this concept can be very approximate, since the drawn curve can touch only a portion of the curves in the family from which it runs a short distance away and some it intersects.

The nature of the dependence and the number of parameters m in describing it are determined from the type of obtained approximate envelope. Then using the method already described in 2.4.5 m points are chosen on this curve and their coordinates x_i, y_i are measured. On the basis of these pairs of coordinates, a system of m equations is drawn up with m unknowns from which are determined the values for the parameters of the approximate envelope.

In determining the type of curve, one should use all the moments of logical analysis into the essence of the process as described in 2.4.4. It must be pointed out that for an envelope such an analysis to a significant degree is difficult due to the fact that its nature is determined by both evolutionary change and revolutionary shifts in the development of the process. However, the demands made upon the extrapolation accuracy in the given instance are significantly less because of the more generalized estimates and the greater, as a rule, lead time in comparison with the simple extrapolation of dynamic series.

In solving the BTS development forecasting problems, the envelope method can be used for forecasting aggregated characteristics of scientific-technical progress in any area of technology. In this instance a certain integral development trend is derived which is invariant to the past and future leaps of a qualitative nature. Examples of such envelopes have been given above.

As one other example one might give the envelope for the speed of generations of computers. Here the C parameter is a vector characterized by the computer circuitry, by the structure and composition of the computer, by the principles of organizing the computer process and by other characteristics which distinguish the different generation machines.

Having constructed a family of curves for the dependences of computer speed for the first, second, third and fourth generations upon output time and having obtained the envelope equation, we will disclose the general trend for the change in speed which does not depend upon the type of circuitry and other indicators determining the computer's generation. In extrapolating this envelope, we obtain forecasts of computer speed for unknown future generations with still unknown operating principles and circuitry.

2.5. Probability and Statistical Methods in Forecast Research

The methods of probability theory and mathematical statistics are used very widely in the various methods and procedures of forecasting. Here it must be pointed out that these areas of mathematics are used both in the methods of processing factographic information and in the processing of data obtained by the expert method. Let us examine certain, in our view, promising applications of probability and statistical methods to forecast research. While the previous section 2.4 examined questions of analyzing development trends for a certain variable in time, in the given section attention has been focused on elucidating the links between two or a range of random variables and the use of the elucidated links for forecasting their values. In terms of the degree of research comprehensiveness, the methods employed in this area can be divided into bivariate and multivariate.

The former involve an examination of paired relationships between variables (paired correlations and regressions) and are aimed in forecast research at solving such problems as establishing the quantitative measure of the closeness of the tie between two random variables, the establishing of the closeness of this tie to a linear one and assessing the reliability and accuracy of forecasts obtained by the extrapolation of a regression dependence.

The multivariate methods of statistical analysis are basically aimed at solving the problem of systems analysis for multivariate stochastic forecasting objects. The aim of such research, as a rule, is to disclose the interrelationships between the variables of the complex, to construct multidimentional functions for the relation of the variables and the isolating of a minimum number of characteristics describing the object with a sufficient degree of accuracy. Here one of the basic tasks is to reduce the size of description for the forecasting object. Thus, statistical methods are basically used to prepare data and to reduce them to a sort suitable for making the forecast. As a rule, after their use one of the extrapolation or interpolation methods is employed for directly obtaining the forecast result.

The mathematics of probability theory and mathematical statistics have been described sufficiently widely in the mathematical literature [8, 21]. The questions of applying these methods to the solving of forecasting problems have also been taken up in the literature [24, 37, 38, 43]. Let us examine the methods of a paired correlation and regression in forecasting, the methods of multivariate statistical analysis and the questions of their application to solving forecasting problems.

2.5.1. Correlations and Regressions in Forecasting Interrelated Random Variables

Let there be a set of values for two random variables $X_1\{x_{1i}\}$ and $X_p\{x_{pi}\}$ for which there is a supposition of the presence of a linear-type relationship with random deviations. Let $\bar{x}_1$ and $\bar{x}_p$--the mean arithmetic values of these variables:

$$\bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{1i}; \quad \bar{x}_p = \frac{1}{n}\sum_{i=1}^{n} x_{pi}. \tag{2.82}$$

The relation between the variables is determined using a paired correlation coefficent

$$r_{1p} = \frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)(x_{pi} - \bar{x}_p)}{S_1 S_p}, \tag{2.83}$$

where S_1, S_p--the standard deviations of x_1 and x_p;

$$S_1 = \sqrt{\frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}{n}} \quad \text{and} \quad S_p = \sqrt{\frac{\sum_{i=1}^{n} (x_{pi} - \bar{x}_p)^2}{n}}.$$

The paired correlation coefficient is employed for assessing the closeness of the relationship if the relationship is a linear one. With a nonlinear form of a relationship a correlation ratio or index is used

$$i_{1p} = \sqrt{1 - \frac{S_{1p}^2}{S_1^2}}, \tag{2.84}$$

where S_{1p}^2--the residual dispersion of x_1 relative to the regression line $x_1 = f(x_p)$;
S_1^2--the full dispersion of x_1 relative to $\bar{x}_1$.

As is seen from (2.84), for determining the correlation index for the actual data it is essential to preselect a regression line and this complicates the procedure of selecting the influencing factors (the method of selecting the form of relationship is given above). In order to avoid intermediate operations related to choosing the regression line, instead of the natural amounts of variables it is possible to use their functions in assuming that the relation between the functions of the studied amounts, for example, $\ln x_1$ and $\ln x_p$ is a linear one.

Subsequently, when the regression equation has been chosen, the regression is tested for linearity. The regression test for linearity is based on the fact that there is a definite relationship between the correlation ratio i and the correlation coefficient r. In the event of a linear correlation association $i^2 = r^2$, in all other instances $i^2 \neq r^2$, and i in terms of its absolute value cannot be less than r. The difference $(i^2 - r^2)$ cannot be negative and if it equals zero, then the regression is a linear one. The regression linearity is assessed from the Fisher z criterion

$$z = \frac{1}{2} \ln \left[\frac{(i^2_{1p} - r^2_{1p})}{(1 - i^2_{1p})} \frac{(n-p)}{(p-2)} \right] \geqslant z_{\mathcal{P}} \qquad (2.85)$$

with the degrees of freedom $\nu_1 = p - 2$ and $\nu_2 = n - p$, where $\mathcal{P}$--fiducial probability.

The significance of the calculated correlation coefficients is assessed from two criteria:

$$t = \frac{r_{1p}}{\sqrt{1 - r^2_{1p}}} \sqrt{n-2} \geqslant t_{\mathcal{P}}; \qquad (2.86)$$

$$z = \frac{1}{2} \ln \frac{1 + r_{1p}}{1 - r_{1p}} \geqslant z_{\mathcal{P}}. \qquad (2.87)$$

A linear regression equation can be written in the form

$$(x_1 - \bar{x}_1) = r_{1p} \frac{S_1}{S_p} (x_p - \bar{x}_p). \qquad (2.88)$$

If r = 0, then a correlation relationship is absent between x_1 and x_p, if r = 1, then then x_1 grows linearly with the growth of x_p, if r = -1, then x_1 diminishes linearly with the growth of x_p. The values $x < |r| < 1$ characterize certain intermediate types of relationship between x_1 and x_p. The coefficient $r_{1p} \frac{S_1}{S_p}$ is termed a linear regression coefficient; it defines the slope of the regression line to the x_p axis.

In forecasting such an extrapolation method is used in the instance that a sufficient amount of statistical data is available for two interrelated variables. One of these variables is chosen as the independent and the second is the dependent (the function). As the independent variable one selects the one from the pair of variables the significance of which can by one means or another be judged for the forecast lead time.

The step paired regression is a particular instance of using a paired regression in forecast research. In this instance by studying a chain of paired variable relationships one arrives at a definition of the forecasted variable. The scheme for applying a step regression thus comes down to the following:

$$x_1 = f_1(x_2),\ x_2 = f_2(x_3),\ x_3 = f_3(x_4) \text{ and so forth.} \qquad (2.89)$$

Using such a scheme, for example, it is possible to investigate the relationship of economic indicators for product costs x_1, labor productivity x_2, the capital-to-labor ratio x_3 and capital investments x_4.

It is essential to bear in mind that the mistakes in determining the final indicator increase rapidly with a rise in the length of the "ladder" of intermediate indicators. In the event of the independent distribution of the random errors u_1, u_2, u_3, ... among themselves with certain constant confidence intervals D_1, D_2, D_3 ... for each step, the probabilities of falling in them will be multiplied in rapidly reducing the significance level of determining the final variable of the "ladder."

In addition to studying the relationships of two random variables in forecasting use is also made of correlation and regression analysis between the values of the same variable. The use of autoregression models for forecasting is based on the assumption that the value of a time series at a moment of time t (with the accuracy up to a random amount) is a linear function of its values at the preceding moments. The modeling process using the autoregression equations consists in isolating from the time series a stationary random component and constructing an autoregression model for it:

$$Z_t = b_1 z_{t-1} + b_2 z_{t-2} + \dots + b_m z_{t-m} + \varepsilon_t. \tag{2.90}$$

The isolating of a stationary random component comes down to excluding the trend from the time series. The simplest and most widespread method of excluding the systematic part of a time series is the method of forming differences of the corresponding order (see Table 2.6).

After the formation of the first order differences the centering of the obtained series Δ^1_t is carried out:

$$\delta_t = \Delta^1_t - \bar{\Delta}^1 \tag{2.91}$$

and the constructing of the autoregression model of the type

$$\delta_t = a_1 \delta_{t-1} + a_2 \delta_{t-2} + \dots + a_m \delta_{t-m} + \eta_t. \tag{2.92}$$

Before beginning to calculate the coefficients for the equation (2.92) it is essential to test the premises of its application in forecasting. The test consists of the two stages examined below.

1. The test of whether the random component is a random amount not dependent on time. This is carried out using the criterion of a series based upon the sampling median. Let the deviations from the trend be calculated: η_1, η_2, ..., η_n. From these an ordered series is formed η^1, η^2, ..., η^n, where n --the least of all deviations. Furthermore, let η_{med}--the median of this ordered series. Then it is possible to form a sequence of the pluses and minuses using the following rule: in place i (i = 1, 2, ..., n) a plus sign is put if observation i in the initial series exceeds the median and a minus sign if it is less than the median; if observation i equals the median, it is dropped. Thus, we obtain a sequence consisting of pluses and minuses the number of which does not exceed n. If the deviations are random

ones then their alternation should also be random. The sequence of running pluses or minuses is termed a series.

Let us calculate the extent $K_{\max(n)}$ of the longest series and the total number of series $V(n)$. In order that the initial series represents a random sample, the extent of the longest series should not be too great while the total number of series should not be too small. The sample is recognized as a random one if the following inequality is met (for a 5 percent confidence level):

$$\begin{cases} K_{\max(n)} < [3{,}3(\lg n + 1)]; \\ V(n) > \left[\frac{1}{2}(n+1) - 1{,}96\sqrt{n-1}\right]. \end{cases} \quad 2.93)$$

If even one of the inequalities is disrupted, then the hypothesis of the random nature of the deviations in the time series levels from the trend is rejected.

2. The test for the hypothesis of stationarity for the random component. The basic condition for the stationarity of a random process is the condition of the dependence of the autocorrelation function solely upon the difference of the independent variables $t_i = t_j = \tau$.

Let us test the hypothesis that the value of the autocorrelation function does not depend upon the choice of the beginning of the observation count but depends just on the amount of the shift of τ. For this for the random component $\eta_t (t = 1, 2, \ldots, n)$ we will find the value of the autocorrelation function r_1^n, r_2^n, ..., r_τ^n (the upper index is the number of observations for which the autocorrelation function is calculated). A formula for calculating an autocorrelation function has the following form:

$$r_{\eta_t}(\tau) = \frac{\sum_{t=1}^{n-\tau} \eta_t \eta_{t+\tau}}{\sum_{t=1}^{n-\tau} \eta_t^2}.$$

Then one of the observation curves is excluded and a new autocorrelation function is calculated r_1^{n-1}, r_2^{n-1}, ..., r_τ^{n-1}. In a similar manner one excludes $P (P = 0, 1, 2, \ldots, p)$ observations and the value of the $(p+1)$ autocorrelation function is calculated. Thus we obtain τ groups of autocorrelation coefficients and each of them will include $(p+1)$ coefficients. For a stationary process, the autocorrelation coefficients included in the same group should be uniform.

The test for uniformity can be carried out in the following manner. For each $r_\tau^{(n-p)}$ included in group τ, we calculate the amount of the z criterion using the formula

$$z_{\tau_p} = \frac{1}{2} \ln \frac{1 + r_\tau^{(n-p)}}{1 - r_\tau^{(n-p)}}. \quad (2.94)$$

Then for this group we calculate

$$\bar{z}_\tau = \frac{\sum_{p=0}^{p} z_{\tau_p}}{p+1}. \qquad (2.95)$$

Mote has proved that the amount

$$\sum_{p=0}^{p} \frac{(z_{\tau_p} - \bar{z}_\tau)^2}{\frac{1}{n} - p - 3} \qquad (2.96)$$

is distributed as χ^2 with p degrees of freedom. In the event that the amount χ^2 calculated using formula (2.96) is less than the tabular value of χ^2 with a set confidence level, then the hypothesis of the uniformity for group τ of autocorrelation coefficients can be accepted. If the uniformity hypothesis is confirmed for all the groups, then it can be accepted that the random component is a random process that is stationary in the broad sense.

If in the test it is established that a time series does not meet even one of these conditions, it is wrong to apply an equation of the type (2.90) for describing it. In this case one must move on to higher order differences.

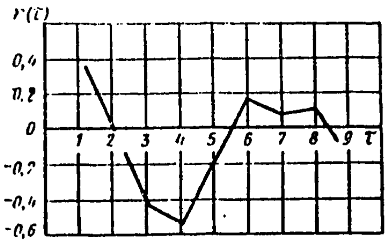


Fig. 2.12. Correlogram for determining the order of an autoregression model

After the necessary tests have been made, it is essential to determine the order of the autoregression model. The first step in selecting the order of the autoaggression model is an analysis of the correlogram. The autoregression process is characterized by the fact that an autocorrelation function is diminishing. A diminishing of the correlogram indicates that the relation with the past is weakening. The autoregression order is determined depending upon in what shifts the autocorrelation function reaches the greatest amount. Let us explain this from a specific example.

From Fig. 2.12 it can be seen that for $\tau > 5$ (τ--the amount of the shift) there is a damping of the correlogram, that is, the relation with the past is weakening. On this basis it can be concluded that for the given example it is not advisable to construct a model above the fourth order. Since an autocorrelation function reaches the greatest amount in the third and fourth shifts, autoregression models of the third and fourth order are constructed.

After first determining the order the parameters of the autoregression model are found. The autoregression parameters can be determined by a double method. The first is the direct application of the least square method. The condition for the minimum dispersion of the deviations in the fixed sample from n observations is written in the form

$$D[\eta_t] = \sum_{t=m+1}^{n} \left(\delta_t - \sum_{j=1}^{m} a_j \delta_{t-j}\right)^2 = \min. \qquad (2.97)$$

This condition leads to a system of normal equations

$$\begin{aligned}
&a_1 \sum_{t=m+1}^{n} \delta_{t-1}^2 + a_2 \sum_{t=m+1}^{n} \delta_{t-1}\delta_{t-2} + \cdots + a_m \sum_{t=m+1}^{n} \delta_{t-1}\delta_{t-m} = \\
&\qquad = \sum_{t=m+1}^{n} \delta_{t-1}\delta_t; \\
&a_1 \sum_{t=m+1}^{n} \delta_{t-1}\delta_{t-2} + a_2 \sum_{t=m+1}^{n} \delta_{t-2}^2 + \cdots + a_m \sum_{t=m+1}^{n} \delta_{t-2}\delta_{t-m} = \\
&\qquad = \sum_{t=m+1}^{n} \delta_{t-2}\delta_t; \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&a_1 \sum_{t=m+1}^{n} \delta_{t-1}\delta_{t-m} + a_2 \sum_{t=m+1}^{n} \delta_{t-2}\delta_{t-m} + \cdots + a_m \sum_{t=m+1}^{n} \delta_{t-m}^2 = \\
&\qquad = \sum_{t=m+1}^{n} \delta_{t-m}\delta_t.
\end{aligned} \tag{2.98}$$

The estimates of the autoregression coefficients can be obtained by another method from a system of Yule-Walker equations

$$\begin{aligned}
&r_1 + a_1 + r_1 a_2 + \cdots + r_{m-1} a_m = 0; \\
&r_2 + r_1 a_1 + a_2 + \cdots + r_{m-2} a_m = 0; \\
&\dots\dots\dots\dots\dots \\
&r_m + r_{m-1} a_1 + r_{m-2} a_2 + \cdots + a_m = 0,
\end{aligned} \tag{2.99}$$

where r_i--the coefficient for the i-order autocorrelation.

For all cases, with the exception of p values, the system of (2.99) is easier to solve than the system (2.98), since there is a more symmetrical form in the construction. As is known the solution to the system (2.99) comes down to recurrent relations

$$a_{ss} = -\frac{r_s + a_{s-1,1} r_{s-1} + a_{s-1,2} r_{s-2} + \cdots + a_{s-1,s-1} r_1}{1 + a_{s-1,1} r_1 + \cdots + a_{s-1,s-1} r_{s-1}} \tag{2.100}$$

$$(s = 1, 2, \dots, m);$$

$$a_{sh} = a_{s-1,h} + a_{ss} a_{s-1,s-h} \; (h = 1, 2, \dots, s-1). \tag{2.101}$$

Here as the initial value one uses $a_{11} = -r_1$. As a result of applying (2.100) and (2.101), we obtain estimates for the coefficients of an autoregression model of the s order, since $a_{s1} = a_1$, $a_{s2} = a_2$, ..., $a_{sm} = a_m$. Having determined the parameters of the autoregression equation, it is now possible to more precisely set the order of the autoregression model. For determining the order of this model the Mann and Wald criteria is employed. It is convenient in the fact that it employs directly calculable and not estimated amounts.

In order to establish whether a sufficiently high degree of approximation has been obtained as a result of the approximation by the model (2.92) of the order m, it is essential to determine the deviations which arise in increasing the autoregression order. For this autoregression models of the q order are constructed, where $m < q < [\frac{n}{2}]$ and the total of the squares of the deviations is determined for those observations for which the value δ_t could be calculated both using the m-order models and the q-order models.

Let Σ_1'--the total of the squares for the remainders for the m-order model (the prime means that not all the remainders are examined but only those for which a q-order model has also been determined); Σ_2--the total of the squares of the remainders for the q-order model. In this instance the amount of the Mann and Wald criteria

$$- (n - q) \ln \Sigma_2/\Sigma_1^* \qquad (2.102)$$

has a distribution χ^2 with $(q - m)$ degrees of freedom. Consequently, having set a given confidence level, it is possible to test the hypothesis that the studied process actually represents an autoregression of the set order. The model can be used for further analysis if satisfactory results have been obtained for the Mann and Wald criterion. Forecasting the value of

Forecasting the value of δ_t for the period $(t + \ell)$ using an autoregression model as carried out in the following manner. Initially the value δ_{t+1}^* is calculated using the formula

$$\delta_{t+1}^* = a_1\delta_t + a_2\delta_{t-1} + \cdots + a_m\delta_{t-m} + \eta_t, \qquad (2.103)$$

then in the expression

$$\delta_{t+2}^* = a_1\delta_{t+1}^* + a_2\delta_t + \cdots + a_m\delta_{t-m} + \eta_t \qquad (2.104)$$

we substitute the calculated value δ_{t+1}^* and determine the value of δ_{t+2}^* and so forth.

In carring out the corresponding transformations, it is essential to move on to a forecast for $(t + 1)$ period of the initial time series.

The transformation must be made using the formulas

$$\Delta_t' = \overline{\Delta}' + \delta_t; \qquad (2.105)$$

$$Y_{t+1} = Y_t' + \Delta_t'. \qquad (2.106)$$

Of substantial significance is an estimate of the error factor of the forecast obtained using the autoregression model. Let us assume that the order of the autoregression model has been set and the model coefficients are known. In this instance the remainders η_t in the autoregression equation are distributed normally with a zero mathematical expectation and a variance σ_η^2. From this it is possible to obtain the probability limits for the predicted value of δ_t:

$$\mathcal{P}\{|\delta_t - \delta_t^*| < t\sigma_\eta\} = \alpha \qquad (2.107)$$

or

$$\mathcal{P}\{\delta_i^* - t\sigma_\eta < \delta_i < \delta_i^* + t\sigma_\eta\} = \alpha, \tag{2.108}$$

where δ_t--the true value of the examined parameter;

δ_t^*--the predicted value of the examined parameter.

Let us now examine the possible methods of assessing the accuracy and reliability of the regression and correlation models. The following statistical characteristics can be employed as indicators of the approximation accuracy.

1. An estimate of the standard error of the regression equation

$$S_{1.f(x)} = \sqrt{\frac{\sum_{i=1}^{n} [y_i - f(x_i)]^2}{n-p}}, \tag{2.109}$$

where n--the number of observations;

p--the number of constants in the regression equation.

2. The mean relative error of the estimate

$$m_\varepsilon = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i - f(x_i)}{y_i} 100\%. \tag{2.110}$$

3. The mean linear error

$$\bar{B} = \frac{\sum_{i=1}^{n} |y_i - f(x_i)|}{\sqrt{n(n-1)}}. \tag{2.111}$$

The characteristics of (2.109)-(2.111) show the degree of approximation of the regression equation to the series of observations of the dependent variable but do not provide any idea about the existence of a relation between the variables from the viewpoint of the sensitivity of the dependent variable to changes in the variable of the independent variable. For this reason the correlation index is a stronger criterion for judging the accuracy of approximation and this is sometimes termed the correlation ratio:

$$J = \sqrt{1 - \frac{S_{1.\bar{f}(x)}^2}{S_1^2}}, \tag{2.112}$$

where S_1^2--the full variance of the dependent variable;

$$S_1^2 = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}; \tag{2.113}$$

here $\bar{y}$--the arithmetic average of the dependent variable calculated from the empirical data of the series.

Since $0 < J < 1$, the closeness of the correlation index to 1 makes it possible to judge simultaneously the quality of approximation and the significance of the relation between the variables. The significance estimate provides an opportunity to make a judgment about the reliability of the selected model. The model's reliability can be set by a significance estimate of the correlation index using the Fisher z criterion:

$$z = \frac{1}{2} \ln\left[\frac{J^2}{(1-J^2)} \frac{(n-p)}{(p-1)}\right] \geqslant z_p, \tag{2.114}$$

where z_p--the tabular value of z with the set fiducial probability [4].

For obtaining sufficiently dependable estimates the level of fiduciary probability must be set not lower than 95 percent, that is $p \geqslant 0.95$. The purpose of estimating the significance of the correlation index is to test the hypothesis of the absence of a relation between the variables. From (2.112) it can be seen that this corresponds to the case where the variance of the data in the series relative to the regression line $S^2_{1.\tilde{f}(x)}$ is close in amount to the variance of the relative mean of the series S^2_1. The question is how significant is the difference of these two indicators, that is, to what degree the substituting of $\bar{y}$ and S^2_1 for $\tilde{f}(x)$ and $S^2_{1.\tilde{f}(x)}$ improves our ideas about changes in the dependent variable. Along with estimating the correlation ratio, the model's reliability can also be judged directly from the variance ratio. For this purpose the Fisher F criterion is used

$$F = \frac{S^2_1}{S^2_{1.\tilde{f}(x)}} \geqslant F_{1-\alpha}, \tag{2.115}$$

where $F_{1-\alpha}$--the tabular value of the Fisher F criterion with the given probability α of rejecting the hypothesis on the equality of the variance ratio [31].

An analysis of the model for sensitivity to changes in the independent variable plays an important role in assessing model reliability. Attempts to judge model sensitivity visually from the position of the data on a graph can lead to serious errors, as the nature of the graph is influenced by the set scale of construction. This is seen from Fig. 2.13, where the same data have been depicted with differing scales for the dependent variable. In Fig. 2.13a the data on the dependent variable are extended along the y axis while in Fig. 2.13b, on the contrary, they are compressed. It is not difficult to realize that it is very hard to judge the existence of trends proceeding solely from graphic presentations.

Thus, the use of the Fisher z and F criteria makes it possible to avoid mistakes in judgments about the real relationship between variables. If several mathematical functions satisfy all the listed requirements equally, preference should be given to the simpler analytical expressions. Here it is essential to consider the possibility of a semantic interpretation of the constants. The latter plays an important role if the obtained model is to be used as the basis for further research on factors which are not considered by the set dependence.

Forecasting from a model selected as a result of the above-described procedure consists in calculating the values of the dependent variable from the values of the independent variable which go beyond the limit of the available aggregate of initial

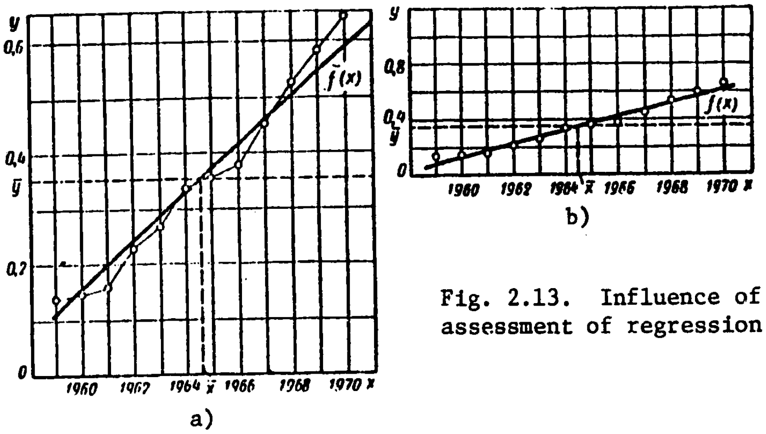


Fig. 2.13. Influence of scale on visual assessment of regression model reliability

data and the establishing of confidence intervals within which one must expect the appearance of the real value of the dependent variable with the determined set probability. Here it is assumed that the prediction error has been distributed normally relative to the regression line and that the mistakes are mutually independent. The latter supposition requires a special check in analyzing dynamic series since the internal series (serial) correlation leads to a significant displacement of the variances in the model parameters and these are incorporated in the general error of the regression equation of [44]. For this reason, before going on to calculate the confidence intervals for the forecasted variable it is essential to test the model for error autocorrelation.

In order to establish the presence of a correlation in a sequential series of values ε_t for a time series, it is essential to use the Darbin-Watson serial correlation criterion [44] which is written in the following manner

$$D = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^2}, \qquad (2.116)$$

where ε_t--the remainder of $(y_t - \tilde{f}(t))$ at the moment of time t;
$\varepsilon_t - \varepsilon_{t-1}$--the righthanded sequential difference of remainders.

The significance of a serial correlation can be assessed from the table which gives the upper and lower limits of the criterion (such a table is found in [44]). If it is established that a correlation of errors occurs, then it is better to use autoregression forecasting methods the contents of which are given below. In the absence of a serial correlation it is possible to move on to calculating the model's general error which is formed from the error of the regression equation and the errors of its coefficients [16, 44]. A diagram for the formation of the general error can be easily traced from the example of a linear model (Fig. 2.14), if the independent variable is expressed through deviations from its mean.

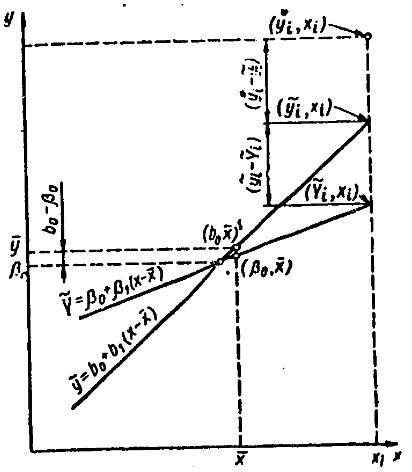


Fig. 2.14. Diagram for the formation of the general error of a linear model (here $\tilde{Y} = \beta_0 + \beta_1(x-\bar{x})$--the regression equation with true coefficients)

As is seen from Fig. 2.14, the general error of the actual amount of y_i^* is formed from the error of the predicted value of y on the regression line with the estimate parameters b_0 and b_1 and from the deviation of the predicted value $\tilde{y}_i$ from the value Y_i on the regression line with true coefficients as caused by the error in the angular coefficient b_1 and by the error of the arithmetic mean $\bar{y}$ caused by the deviation of b_0 from β_0. In knowing the amount of each of the particular errors, it is possible to determine the general error using the formula

$$S_{\{y_i^*-\tilde{y}_i\}} = \sqrt{S_{1.f(x)}^2 + \frac{S_{1.f(x)}^2}{n} + \frac{S_{1.f(x)}^2}{S_x^2\sqrt{n}}(x_i-\bar{x})^2}$$

or

$$S_{\{y_i^*-\tilde{y}_i\}} = S_{1.f(x)}\sqrt{1+\frac{1}{n}+\frac{1}{\sqrt{n}}\frac{(x_i-\bar{x})^2}{S_x^2}}. \tag{2.117}$$

The confidence interval for the individual prediction of the value of y_i^* can be presented in the form of an inequality

$$\tilde{y}_i - t_pS_{\{y_i^*-\tilde{y}_i\}} < y_i^* < \tilde{y}_i + t_pS_{\{y_i^*-\tilde{y}_i\}}, \tag{2.118}$$

where $\tilde{y}_i$--the point estimate of the dependence of the variable on the estimate regression line;

t_p--the value of the Student-Gosseth t-criterion for the set level of fiduciary probability with n degrees of freedom (n--the number of terms from the dynamic series [16, 44]).

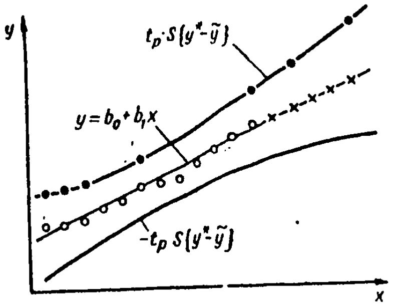


Fig. 2.15. Confidence interval for individual value of dependent variable estimated by linear model

An analysis of the expression (2.117) indicates that the confidence interval changes along the regression line from the minimum value with $x_i = \bar{x}$, when the third term of the expression under the radical turns to zero, in maximally widening toward the end values of the independent variable, as is shown in Fig. 2.15. The point y_i^* on the graph is the extrapolation value of the dependent variable for x_{n+1} value of the dependent variable which for a dynamic series is equivalent to (T+1) year.

The possibility of employing confidence intervals calculated from the data of observations to estimate extrapolation reliability has been examined in detail in [8]. However it is not difficult to see that this reliability rapidly diminishes with an increase in the forecast lead period and with sufficiently high for trend extrapolation loses its practical significance. For this reason, the above-described forecasting method should be employed for short-term forecasting. However it must be kept in mind that even a carefully selected model can be inadequate to a complete dynamic series although it is the best for a part of this series. This is the basic difficulty in forecasting using regression equations. They bear too strong an impression from the previous history of the process, in describing its patterns as an average, that is, equally good or bad for all development phases of the object.

2.5.2. Multivariate Statistical Analysis

In a majority of real studies of the BTS, the forecasting object is a multidimensional complex described by an aggregate of interrelated variables. The difficulty of systems analysis and forecasting for such a complex increases sharply with a rise in the number of variables comprising it. On the other hand, a study of the development of an individually taken variable from the complex or of the interaction of the pair of variables is always fraught with the danger of obtaining erroneous forecasts due to the overlooking of the effect of other variables. For this reason the forecaster is confronted with the question of choosing the research method which will provide a compromise between the complexity of realizing it and the reliability of the obtained results.

The multidimensionality of the relationships and the random nature of a number of variables in the forecasting object necessitate an examination of it as a stochastic object and, consequently, the use of the statistical and probability methods of multivariate analysis. Here the following approaches can be used:

1) Multivariate regression and correlation analysis aimed at constructing the surfaces (planes) of multiple regression for the studied characteristic with the subsequent interpolation or extrapolation of its value;

2) Statistical analysis of the internal structure of the complex aimed at disclosing the main (determining, leading) variables which define the development of the related secondary variables; this, in essence, is the path of minimizing the dimensionality in describing the complex;

3) Factor analysis aimed at disclosing new random variables, each of which is a generalization of a whole series of initial variables in the complex; this approach, like the preceding one, is aimed at minimizing the dimensionality of the object's description but not by discarding unimportant ones and examining the basic ones but rather by increasing the degree of research aggregation.

Multiple regression and correlation analysis is applicable in analyzing the following propositions [11, 22, 43]: the random value distribution laws are close to the normal; the variants of the dependent variable are equal with different values of the independent variables; the independent variables are measured without errors; individual measurements for all amounts are considered stochastically independent, that is, the absence of an autocorrelation is assumed in their change.

There are the following differences in the premises for employing these two methods: regression analysis presupposes the absence of a correlation between the independent variables, while correlation analysis assumes the presence of such a reciprocal correlation; regression analysis is suitable for describing objects of both a linear and nonlinear character while correlation analysis is suitable only for linear objects, since for nonlinear ones there is no sense in the concept of a multivariate normal distribution of the variables and this is an essential condition for its realization.

In solving practical problems, the strict observance of all the listed conditions, as a rule, is impossible. However both methods are applicable considering that the deviations from the above-listed conditions reduce the reliability of results. This must always be kept in mind in assessing the accuracy and reliability of the forecasts obtained by regression and correlation methods.

In the general instance a multiple regression equation represents a dependence between one and several variables. For modeling, for example, the cost estimates of the BTS, linear and nonlinear equations are used for the independent variables and linear ones for parameters. For example, the equation

$$\tilde{X}_1 = b_1 + b_2x_2 + b_3x_3 + \cdots + b_px_p \qquad (2.119)$$

is a linear one for the independent variables and parameters. The equation

$$\tilde{X}_1 = b_1x_2^{b_2}x_3^{b_3} \ \ldots \ x_p^{b_p} \qquad (2.120)$$

is nonlinear for the independent variables but linear for the parameters, as it can be reduced to a linear type by replacing the variables with their standard functions, for example,

$$\log \tilde{X}_1 = \log b_1 + b_2 \log x_2 + b_3 \log x_3 + \cdots + b_p \log x_p.$$

As is seen, the equation (1.20), in contrast to a linear equation, expresses a linear relationship between the logarithms of the variables, that is, it is logarithmically linear. Sometimes for forecast modeling semilogarithmic linear equations are used. For example, the equation

$$\tilde{X}_1 = e^{b_1 + b_2x_2 + b_3x_3 + \cdots + b_px_p} \qquad (2.121)$$

is semilogarithmically linear, as it can be reduced to the linear form:

$$\ln \tilde{X}_1 = b_1 + b_2x_2 + b_3x_3 + \cdots + b_px_p.$$

One other variety of parametric-linear models is the equations which are linear relative to the double logarithm of the independent variable and the logarithms of the dependent variables:

$$\tilde{X}_1 = e^{b_1x_2^{b_2}x_3^{b_3} \cdots x_p^{b_p}}, \qquad (2.122)$$

which has the following linear form:

$$\ln(\ln \tilde{X}_1) = \ln b_1 + b_2 \ln x_2 + b_3 \ln x_3 + \cdots + b_p \ln x_p.$$

As in the case of constructing regression equations with two variables, modeling with multiple regression equations necessitates the choice of a form of the relationship between the variables. The choice of the form of relationship between the variables of a multiple regression equation is complicated by the fact that preliminary analysis using graphic constructs, as is employed in selecting two-variable functions, is not suitable for an analysis of multiple relations. Any of the equations (2.119)-(2.122) expresses the relation between two variables (dependent and independent) under the condition that all the remaining variables in the equation assume certain fixed values. For this reason graphic analysis of the type of dependence cannot be carried out before the parameters of the regression equation b_1, b_2, ..., b_p are found.

Thus, graphic analysis of the form of relation requires the preliminary positing of a hypothesis on the form of the multiple regression equation and the calculating of its parameters. Regardless of this, in certain situations graphic analysis can be a useful auxiliary tool in constructing such models.

Graphic analysis of the type of dependence between the dependent and each of the independent variables in the model is carried out by the sequential elimination of the influence from other variables and this is achieved by substituting into the equation the values of these variables which are numerically equal to their arithmetic means. For example, in a case with three independent variables, a model suitable for analyzing the type of dependence between the variables x_1 and x_2 assumes the following form:

$$\tilde{X}_1 = b_1 x_2^{b_2} \bar{x}_3^{b_3} \bar{x}_4^{b_4}, \tag{2.123}$$

where $\bar{x}_3$ and $\bar{x}_4$--arithmetic means of eliminated variables.

After transformation we arrive at a simple regression equation with two variables:

$$\tilde{X}_1 = b_1' x_2^{b_2}, \; b_1' = b_1 \bar{x}_3^{b_3} \bar{x}_4^{b_4}. \tag{2.124}$$

In an analysis of dependence graphs it must be remembered that each of the amounts of the last product contains a certain error since the calculated values b_1, $\bar{x}_3$ and $\bar{x}_4$ are only estimates of their true values in the general aggregate. For this reason the regression line described by the equation (2.124) can be shifted along the vertical axis and the actual points can be above or below the regression line. However such a circumstance does not influence the nature of the trend.

The effectiveness of aposteriori graphic analysis depends largely upon the number of independent variables, their variance and reciprocal influence. With a large number of independent variables with significant variances and complex internal relationships, most often the placement of factual data in a rectangular system of coordinates does not make it possible to provide any reliable judgment on the type of dependence using geometric constructs. This is quite natural since the multiple relations are described not by lines but by complex spatial figures. In this regard, if there is no definite informal feature for which preference should be given to one of the models beforehand, the choice of the form of relation can be made only on the basis of the regular estimate criteria.

The parameters of a parameter-linear multiple regression equation are set by the least square method. Here there is a minimization of the total of the squares of the actual data deviations from the regression surface described by the given equation:

$$\sum_{i=1}^{n} [x_{1i} - \tilde{x}_{1i}(\bar{X}, \bar{B})]^2 \to \min, \tag{2.125}$$

where $\bar{X}$--the vector of independent variables $(x_2, x_3, \ldots, x_p)$;
$\bar{B}$--the vector of the regression equation parameters $(b_1, b_2, \ldots, b_p)$.

In differentiating (2.125) for each parameter and setting the particular derivatives at zero, we obtain a system of linear equations:

$$\begin{aligned}
\sum x_1 &= nb_1 + b_2\sum x_2 + b_3\sum x_3 + \cdots + b_p\sum x_p;\\
\sum x_1x_2 &= b_1\sum x_2 + b_2\sum x_2^2 + b_3\sum x_2x_3 + \cdots + b_p\sum x_2x_p;\\
\sum x_1x_3 &= b_1\sum x_3 + b_2\sum x_2x_3 + b_3\sum x_3^2 + \cdots + b_p\sum x_3x_p;\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sum x_1x_p &= b_1\sum x_p + b_2\sum x_2x_p + \cdots + b_{p-1}\sum x_{p-1}x_p + b_p\sum x_p^2,
\end{aligned} \tag{2.126}$$

where the totaling is carried out for the entire aggregate of initial data $(i = \overline{1, n})$.

The solving of the system of (2.126) makes it possible to find the parameters of the regression equation $b_1, b_2, \ldots, b_p$ which are optimum for the criterion of (2.125). For calculating the parameters of nonlinear models the latter are linearized (see above). In this case the variables are replaced by their logarithmic functions.

The variables of multiple regression equations can be set in a natural scale, in the form of simple and standardized deviations from their means. For an analysis of the relations between the variables of the greatest interest are the models constructed from the standardized deviations of the variables (standardized variables). The unit of measurement for the degree of reciprocal influence of the variables is their mean square deviation and the formula for converting to a standardized scale is:

$$z_j = \frac{x_j - \bar{x}_j}{\sigma_j}, \quad j = \overline{1, p}, \tag{2.127}$$

where z_j--the new standardized variable;
x_j--the variable in a natural scale;
$\bar{x}_j$--the arithmetic average of variable j;
σ_j--the mean square deviation for variable j;

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij}-\bar{x}_j)^2}{n-1}}. \qquad (2.128)$$

The multiple regression equation expressed in a standardized scale has the form:

$$\tilde{Z}_1 = \beta_2 z_2 + \beta_3 z_3 + \cdots + \beta_p z_p, \qquad (2.129)$$

where $\beta_2, \beta_3, \ldots, \beta_p$--standardized multiple regression coefficients which are determined from the system of linear equations

$$\begin{aligned} \beta_2 &= r_{12} - (\beta_3 r_{23} + \beta_4 r_{24} + \cdots + \beta_p r_{2p}); \\ \beta_3 &= r_{13} - (\beta_2 r_{32} + \beta_4 r_{34} + \cdots + \beta_p r_{3p}); \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_p &= r_{1p} - (\beta_2 r_{p2} + \beta_3 r_{p3} + \cdots + \beta_{p-1} r_{p\,(p-1)}), \end{aligned} \qquad (2.130)$$

where r_{ij}--coefficients for the paired correlation between variables.

The system of (2.130) is obtained as a result of applying the least square method to the standardized variables. The standardized multiple regression coefficients indicate to what part of σ_1 there will be a change in the average value of the dependent variable if the corresponding independent variable changes by one mean square error (σ_j, $j = \overline{2, p}$), while the other variables remain constant. Since all the variables are expressed in comparable units of measurement, the regression coefficients indicate the comparative strength of the influence of each variable on a change in the dependent variable. If the equation has been selected sufficiently correctly, the general influence should be close to one, that is, $\sum_{j=2}^{p} |\beta_j| \approx 1$. Thus, the expressing of multiple regression equations by standardized variables provides a valuable tool for analyzing the sensitivity of the estimates to a change in the independent variables.

The conversion of the equation (2.129) to a natural scale is carried out using the formulas

$$b_j = \beta_j \frac{\sigma_1}{\sigma_j}, \quad j = \overline{2, p}; \qquad (2.131)$$

$$b_1 = \bar{x}_1 - (b_2 \bar{x}_2 + b_3 \bar{x}_3 + \cdots + b_p \bar{x}_p), \qquad (2.132)$$

where $\bar{x}_1$--the arithmetic mean of the dependent variable.

The constructing of equations in a standardized scale involves carrying out additional calculations and the complicating of the calculation procedure for the model parameters. For this reason, with the lack of a well organized program making it possible to carry out the entire calculation procedure on a computer, it is simpler to use the method of calculating the variables on a conventional scale. In this instance, an analysis of the models for sensitivity can be carried out using a simplified scheme with a multiple determination coefficient

$$R'_{1.23\ldots p}=\frac{b_2\sum x_1x_2+b_3\sum x_1x_3+\cdots+b_p\sum x_1x_p}{\sum x_1^2}, \tag{2.133}$$

where x_1, x_2, ..., x_p--deviations of the variables from their averages.

Having divided term by term the righthand portion by $\sum x_1^2$, it is possible to measure the degree of influence of each indicator for independent variable of the model on the estimate of the function as expressed in fractions of the overall variation.

An estimate of the approximation accuracy and model reliability is carried out using two basic criteria:

The multiple correlation coefficient

$$R_{1.23\ldots p} = \sqrt{1-(1-r^2_{1.2})(1-r^2_{13.2})\ldots(1-r^2_{1p.23\ldots(p-1)}}, \tag{2.134}$$

where $r_{13.2}$, ..., $r_{1p.23\ldots(p-1)}$--partial correlation coefficients between the dependent and independent variables;

The multiple correlation index

$$J_{1.23\ldots p} = \sqrt{1-\frac{S^2_{1.f(23\ldots p)}}{S_1^2}}, \tag{2.135}$$

where $S^2_{1.f(23\ldots p)}$--estimate for the variance of the actual data on the relative regression surface;

S_1^2--estimate for the full variance of the actual values for the dependent variable relative to its mean.

An estimate of the variances is made using the formulas:

$$S^2_{1.f(23\ldots p)} = \frac{\sum_{i=1}^{n}(x_{1i}-\tilde{x}_{1i})^2}{n-p}; \tag{2.136}$$

$$S_1^2 = \frac{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2}{n-1}. \tag{2.137}$$

The multiple correlation coefficients and index are always positive and assume any values between 0 and 1. If the corresponding index is equal to or close to zero this shows the absence of a relation between the independent and dependent variables while the closeness of the indicator to one shows a strong relation. From (2.135) one can see that the correlation index $J_{1.23\ldots p} \approx 0$, if $S_{1.f(23\ldots p)} \approx S_1 J_{1.23\ldots p} \approx 1$, when $S_{1.f(23\ldots p)} \approx 0$.

The latter means that virtually all the actual values of the dependent variable lie on the regression surface or in any event are located in the immediate proximity of it.

The sense of the ratio of the two variances in (2.135) comes down to the fact that by using it it is possible to elucidate how much better our estimates are described by a regression equation in comparison with the average value of the dependent variable. In considering that the observed difference in the variances can be the result of a random variation of the data, it is essential to estimate the significance of the established relations. This estimate can be carried out using the Fisher z-criterion

$$z = \frac{1}{2} \ln \frac{J^2}{1 - J^2} \frac{n - p}{p - 1} \geqslant z_{\wp} \tag{2.138}$$

with the degrees of freedom $\nu_1 = (p - 1)$ and $\nu_2 = (n - p)$.

The same criterion is employed for assessing the significance of the multiple correlation coefficient. The estimate of the significance of the multiple correlation coefficients and index makes it possible to judge the reliability of the model. The The same indicators are employed for estimating the approximation as the multiple correlation coefficient makes it possible to choose the better linear model while the multiple correlation index does the same for the best nonlinear one. A comparison of them can finally establish the form of the relation. For this it is essential to assess the linearity of the relation using the Fisher z-criterion:

$$z = \frac{1}{2} \ln \frac{J^2_{1.23\ldots p} - R^2_{1.23\ldots p}}{1 - J^2_{1.23\ldots p}} \frac{n - p}{p - 2} \leqslant z_{\wp} \tag{2.139}$$

With degrees of freedom $\nu_1 = (p - 2)$ and $\nu_2 = (n - p)$. If the condition of (2.139) is satisfied, the linearity of the relation is considered established and preference must be given to a linear model.

The accuracy of the forecast estimates is characterized by the square of the error for the individual prediction which is calculated using the formula

$$\begin{gathered} S^2_{x_1 - \tilde{x}1.f(23\ldots p)} = S^2_{1.f(23\ldots p)} \Big(1 + \frac{1}{n} + c_{22}x_2^2 + c_{33}x_3^2 + \\ + \cdots + c_{pp}x_p^2 + 2c_{21}x_2x_3 + 2c_{21}x_2x_4 + \cdots + 2c_{jp}x_jx_p \Big), \end{gathered} \tag{2.140}$$

where $S_{1.f(23\ldots p)}$--estimate of the standard error relative to the regression surface;

$x_2, x_3, \ldots, x_p$--independent variables expressed in the form of deviations from the corresponding averages;

n--the volume of the initial statistical aggregate;

$c_{22}, c_{33}, \ldots, c_{pp}, c_{23}, c_{24}, \ldots, c_{jp}$--constants determined from the system of equations in [16]:

$$\begin{cases} (\sum x_2^2)\,c_{22} + (\sum x_2x_3)\,c_{23} + \cdots + (\sum x_2x_p)\,c_{2p} = 1; \\ (\sum x_2x_3)\,c_{22} + (\sum x_3^2)\,c_{23} + \cdots + (\sum x_3x_p)\,c_{2p} = 0; \\ \cdot\;\cdot \\ \cdot\;\cdot \\ (\sum x_2x_p)\,c_{22} + (\sum x_3x_p)\,c_{23} + \cdots + (\sum x_p^2)\,c_{2p} = 0; \end{cases}$$

$$\begin{cases} (\sum x_2^2)\,c_{32} + (\sum x_2x_3)\,c_{33} + \cdots + (\sum x_2x_p)\,c_{3p} = 1; \\ (\sum x_2x_3)\,c_{32} + (\sum x_3^2)\,c_{33} + \cdots + (\sum x_3x_p)\,c_{3p} = 0; \\ \cdot\;\cdot \\ \cdot\;\cdot \\ (\sum x_2x_p)\,c_{32} + (\sum x_3x_p)\,c_{33} + \cdots + (\sum x_p^2)\,c_{3p} = 0; \end{cases} \tag{2.141}$$

$$\cdot\;\cdot$$

$$\begin{cases} \cdot\;\cdot \\ (\sum x_2^2)\,c_{p2} + (\sum x_2x_3)\,c_{p3} + \cdots + (\sum x_2x_p)\,c_{pp} = 0; \\ (\sum x_2x_3)\,c_{p2} + (\sum x_3^2)\,c_{p3} + \cdots + (\sum x_3x_p)\,c_{pp} = 0; \\ \cdot\;\cdot \\ (\sum x_2x_p)\,c_{p2} + (\sum x_3x_p)\,c_{p3} + \cdots + (\sum x_p^2)\,c_{pp} = 1. \end{cases}$$

The deviations of the independent variables from the averages for the nonlinear regression equations are calculated from the formula

$$x_j = x_j^{b_j} - \bar{x}_j^{b_j}, \tag{2.142}$$

where $\bar{x}_j$--arithmetic average of independent variable in initial statistical aggregate.

Since it is assumed that the error of an individual prediction is subordinate to a normal distribution law, the confidence interval within which one can expect the forecasted amount of the dependent variable with the given probability can be determined by the expression

$$\tilde{x}_1 - t_{\mathscr{P}} S_{x_1^{\bullet} - \tilde{x}_{1.j\,(23\ldots p)}} \leqslant x_1^{\bullet} \leqslant \tilde{x}_1 + t_{\mathscr{P}} S_{x_1^{\bullet} - \tilde{x}_{1.j\,(23\ldots p)}}, \tag{2.143}$$

where $\tilde{x}_1$--the predicted value of the dependent variable for the multiple regression equation.

Forecasting from models which represent a linear or nonlinear combination of time or statistical dependences involves the necessity of calculating the total (integral) forecast error. The total forecast error is defined for a linear combination of dependences as

$$S_{\Sigma} = \sqrt{\sum_{\kappa} S_{\kappa}^2}, \tag{2.144}$$

and for a nonlinear combination using the formula

$$S_{\Sigma} = \sqrt{\sum_{\kappa} \left(\frac{\partial F}{\partial f_{\kappa}} \right) S_{\kappa}^{2}}\,, \qquad (2.145)$$

where S_k--the individual prediction error for dependence k;

$\frac{\partial F}{\partial f_k}$--the partial derivative of a composite model for dependence k.

Let us examine the question of using mathematical statistics to reduce the dimensionality for describing a stochastic object.

The correlation pleiad method is designed to disclose variable groups in a complex the relation between which is the closest to linear. After discovering such groups, all the variables of the group (pleiad) in the research are replaced by one determining variable. As a result the initial dimensionality of description is substantially reduced with small losses in accuracy which are determined by the accepted level of the correlation coefficients.

In the process of such research a correlation matrix is constructed for all the variables of the studied object $D = |r_{ij}|^{n}$.

After this a number of threshold levels are set for constructing the correlation pleiads. For example,

$$\begin{aligned} &1)\ 1 < r_{ij} \leqslant 0{,}9; \\ &2)\ 0{,}9 < r_{ij} \leqslant 0{,}8; \\ &3)\ 0{,}8 < r_{ij} \leqslant 0{,}7. \end{aligned} \qquad (2.146)$$

Then, from the correlation matrix one selects the relations corresponding to each interval. For analysis these can be conveniently presented in the form of a graphic interpretation, an example of which is shown in Fig. 2.16. In this drawing around the circumference lie sequential points corresponding to the variables $x_1, x_2, \ldots, x_n$. In examining the D matrix, one connects with straight lines those variables on the circumference the correlation coefficient between which fell within the set interval 1, 2, or 3 as determined by (2.146).

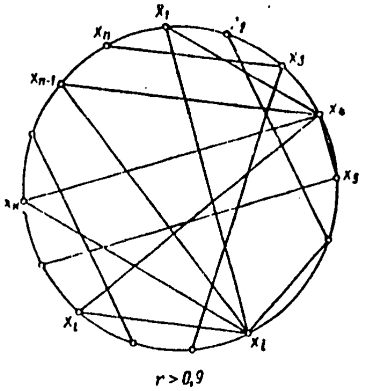


Fig. 2.16. Illustration of the correlation pleiad method

As a rule, the picture obtained as a result will contain a number of clusters analogous to the clusters in x_4 and x_i in Fig. 2.16. In examining such a picture, it is easy to isolate the vertices of the clusters which collect a large number of relations. For level 1 of (2.146) it can be said from

Fig. 2.16 that the variables x_1, x_5, x_ℓ, x_k and x_{n-1} are related virtually functionally (with $r > 0.9$) with the variable x_4; for this reason the analysis and forecast can be made only for x_4 and the obtained results extended to all the variables of the cluster through the formulas of functional ratios. An analogous conclusion can be drawn for the other peak of the cluster x_i in Fig. 2.16.

The construction of several such circles of correlation pleiads for the various levels of correlation coefficients makes it possible to analyze the internal relationships of all the variables in the complex and select for examination a minimum number of them as determined by the given accuracy of the research and the nature of the D matrix.

Another approach to solving the problem of minimizing the system's dimensionality for an object's variables is given in [5, 39] from the standpoint of information theory and pattern recognition theory. From these positions the problem of forecasting BTS development can be represented as the task of recognizing the future state of an object from the results of forecasting the values of a set of individual variables which comprise its description.

The problem of minimizing the demensionality of a description is carried out in the following manner. On the basis of the retrospective values of the variables, from them a minimum number is chosen making it possible with the set reliability to distinguish the states of the forecasting object which are of interest to us. The basic propositions and assumptions in the given examination come down to the following. For the studied object we know a finite, denumerable set $A = \{A_1, A_2, \ldots, A_m\}$ of possible states called classes. The object is described by the set of $X = \{x_1, x_2, \ldots, x_N\}$ variables, each of which can assume a finite number of values. It is assumed that all the variables of x_n are statistically independent of one another, as otherwise the problem becomes particularly cumbersome for practical use.

According to Shannon the information content of a certain variable relative to a set of classes A can be defined as the difference of the initial entropy in the system and the entropy of a solution for the variable x_n

$$J_n = H(A) - H(A/x_n). \qquad (2.147)$$

Let the variable x_n assume T discrete values x_{nj} ($j = 1, 2, \ldots, T$). Then the entropy of the solution with value j of variable x_n can be written:

$$H_j(A/x_n) = -\sum_{m=1}^{M} p(A_m/x_{nj}) \log p(A_m/x_{nj}). \qquad (2.148)$$

According to the Bayes formula

$$p(A_m/x_{nj}) = \frac{p(A_m)\,p(x_{nj}/A_m)}{p(x_{nj})} = \frac{p(A_m)\,p(x_{nj}/A_m)}{\sum_{m=1}^{M} p(A_m)\,p(x_{nj}/A_m)}. \qquad (2.149)$$

In substituting this expression in (2.148), we obtain

$$H_j(A/x_n) = -\frac{1}{p(x_{nj})} \times$$
$$\times \sum p(A_m)\,p(x_{nj}/A_m)\log\frac{p(A_m)\,p(x_{nj}/A_m)}{\sum\limits_{m=1}^{M} p(A_m)\,p(x_{nj}/A_m)} =$$
$$= -\frac{1}{p(x_{nj})}\left[\sum_{m=1}^{M} p(A_m)\,p(x_{nj}/A_m)\log p(A_m)\,p(x_{nj}/A_m) - \right.$$
$$\left. - \sum_{m=1}^{M} p(A_m)\,p(x_{nj}/A_m)\log\sum_{m=1}^{M} p(A_m)\,p(x_{nj}/A_m)\right]. \qquad (2.150)$$

In these formulas: $p(x_{nj})$--the probability of the appearance of gradation j of variable n for all classes of A_m; $p(x_{nj}/A_m)$--the probability of the appearance of value j of variable n for class m.

For obtaining the full entropy of the solution for the variable x_n it is essential to total $H_j(A/x_n)$ for all gradations of j with weights proportional to the probabilities of the appearance of each gradation, that is, $p(x_{nj})$. As a result, we obtain

$$H(A/x_n) = \sum_{j=1}^{T} p(x_{nj})\,H_j(A/x_n) =$$
$$= -\sum_{j=1}^{T}\sum_{m=1}^{M} p(A_m, x_{nj})\log p(A_m, x_{nj}) +$$
$$+ \sum_{j=1}^{T}\sum_{m=1}^{M} p(A_m, x_{nj})\log\sum_{m=1}^{M} p(A_m, x_{nj}).$$

In substituting the obtained expression and the initial entropy $H(A) = -\sum\limits_{m=1}^{M} p(A_m)\log p(A_m)$ in (2.147), we obtain the final expression for the information content of variable n:

$$J_n = -\sum_{m=1}^{M} p(A_m)\log p(A_m) + \sum_{j=1}^{T}\sum_{m=1}^{M} p(A_m, x_{nj})\log p(A_m, x_{nj}) -$$
$$- \sum_{j=1}^{T}\sum_{m=1}^{M} p(A_m, x_{nj})\log\sum_{m=1}^{M} p(A_m, x_{nj}), \qquad (2.151)$$

where $p(A_m, x_{nj})$--the joint distribution of probabilities for the values of x_{nj} for the class A_m.

The expression (2.151) is the basic working formula for calculating the information content of variables. Having determined quantitatively the values for the information content of the variables, it is possible to rank them according to the decline in the values of J_n and determine that minimum number of variables which is essential for recognizing the state of the object with the set degree of reliability.

Let the solution be made by calculating the aposteriori probabilities of the possible states for the values of the variables x_n determined relative to a certain moment of time in the future. Here it is advisable to consider the variables with the highest information content out of the entire set. Thus, the route of examining the variables is set by their ranking number from the decline of J_n. The length of the route, that is, the necessary number of variables in the examination, is determined by the given threshold value for the aposteriori probability of the class of states P_p.

With the least favorable distribution of the aposteriori probabilities of the hypotheses we have

$$P_p; \quad \frac{1-P_p}{M-1}; \quad \frac{1-P_p}{M-1}; \ldots; \quad \frac{1-P_p}{M-1}; \tag{2.152}$$

with the most favorable

$$P_p, \; 1-P_p, \; 0, \; 0, \; \ldots, \; 0. \tag{2.153}$$

The last coordinate of the minimal route is determined by the Fano inequality:

$$LH(A) - H(A/x_1) - H(A/x_2) - \cdots - H(A/x_L) \geqslant H(A) - H_p, \tag{2.154}$$

where L--the number of steps in the minimum route;
H_p--the entropy of the solution determined from the distribution of (2.152).

The above-examined method of minimizing dimensionality in a description of a stochastic forecasting object can be realized in those instances when the problem is posed from the above-described positions of pattern recognition theory and the retrospective analysis provides sufficient statistics of the variables and the states for calculating conditional probabilities.

In the event that in examining the minimal route none of the hypotheses has reached the given threshold P_p, it is possible to use the following variables of the full route. Then two variations are possible: either at a certain step the probability of one of the hypotheses m exceeds P_p and then one can speak about the forecast of the object's state close to m, or, if none of the hypotheses reaches the set probability threshold, then one can speak about the appearance of a fundamentally new state of the object in the future. In lowering the threshold P_p, it is possible to determine to which of the known this new state is closest.

The designated procedures for determining the information content of variables, for seeking out the minimal route and for classifying states in examining it can be successfully carried out on a computer. Experiments in identifying classes of technical systems have provided positive results.

Furthermore, a method has been provided of multivariate statistical analysis making it possible to minimize the description of a stochastic object not by discarding inessential variables but by disclosing certain general characteristics of their aggregate.

2.5.3. Certain Information on Factor Analysis

Factor analysis in its present form represents a certain area of mathematical statistics [26, 42]. However, its appearance at the beginning of the present century is usually linked with the names of the psychologists C. Spearman, S. Barthow, L. Thurstone and other. Its initial aim was to construct mathematical models for human capabilities and conduct. Here the results were based on various psychological and physical tests and at the output certain general indicators or factors were formed. In this area factor analysis is successfully employed at present, however in recent decades it has spread actively into many other areas such as sociology, economics, geology, meteorology, engineering and so forth. The work of H. Harman [42] gives over 200 publications on factor analysis and its use in over a score scientific areas. He also points to the great diversity of factor analysis methods and their modifications which are presently known. Let us examine the essence of one of them.

Let X--an n-dimensional random vector representing a random sample of measurements for an aggregate of interrelated parameters x_i; F--a k-dimensional vector the components of which are directly unobservable variables (the factors F_j); $\overline{X}$--the mathematical expectation of the fector X; U--the vector of the total of the unobservable errors and specific factors. According to the basic assumption of multifactor analysis, each specific measurement of the X-x_i vector can be viewed as the total of the effects of a certain small number of group factors f_j (taken with certain weights a_{ij}), the specific factor s_i effecting only the given variable and the measurement errors e_i. Since s_i and e_i are indistinguishable in factor analysis, they are ordinarily viewed as the total $u_i = s_i + e_i$.

Further, let A--the matrix of the order $n+k(n>k)$, the elements of which are factor weights a_{ij} determining the load of variable i on factor j; m--the number of observations on vector X for which the estimate is made. Let us write the basic idea of factor analysis in a matrix form:

$$X = AF + \overline{X} + U. \quad (2.155)$$

For the sake of simplicity let us set all the averages at zero: $\overline{X} = 0$, that is, we will further view the unbiased distributions x_i. Let us designate the product AF by Q, then

$$X = Q + U, \quad (2.156)$$

where Q--is usually called the general part, and
U--the specific part of X.

It is assumed that U does not depend on Q and all the $u_i (i = 1, 2, \ldots, n)$ are not intercorrelated. Here the matrix $M(QU') = 0$, the matrix $M(UU')$ is diagonal (M--the operator of the mathematical expectation; U'--the transposed matrix U). Then

$$M(XX') = M[(Q+U)(Q'+U')] = M(QQ')+M(UU')+M(QU')+M(UQ') = M(QQ')+M(UU'). \quad (2.157)$$

If we normalize the X vector for the values of the standard deviations $\sigma_i (z_{ip} = x_{ip}/\sigma_i)$ where x_{ip}--a component of the X vector; p--the ordinal number of a single observation of vector X), then M(ZZ') = R is a correlation matrix. In accord with (2.157), this can be shown in the form

$$R = R_0 + U^2 = R_0 + I + H^2, \qquad (2.158)$$

where R--the initial matrix with units on the main diagonal; R_0--the so-called reduced matrix; U^2--a diagonal matrix from the squares of the total specific factor weights and errors; H^2--a diagonal matrix of the so-called communalities; I--a unit matrix.

In factor analysis usually the R matrix determined from (2.158) is subject to factortorization. Here the reduced matrix is approximated by the product

$$R_0 = A_0 A_0', \qquad (2.159)$$

where A_0--is taken as the matrix of the factor weights of the order $(n \times k)$; A_0' --the matrix transposed in relation to A_0.

The factors $f_1, f_2, \ldots, f_k$ are assumed to be uncorrected. In this instance, the factor weights can be viewed as coefficients in a linear regression equation for estimating the variables by the factors.

If we disregard U in (2.156), then A_0 coincides with A and R_0 coincides with R; consequently, the factorization will be closer to the original the closer the matrix of communalities H^2 is to one.

In estimating A_0, usually the main component method is employed, the idea of which is the following. Since R_0 is a real symmetrical matrix, then by the orthogonal similarity transformation, it can be reduced to a diagonal type

$$B^{-1} R_0 B = L, \qquad (2.160)$$

hence $R_0 = BLB'$, or due to the orthogonality of B

$$R_0 = BLB^{-1}; \qquad (2.161)$$

here L--the diagonal matrix comprised of the characteristic roots of R_0 considering their multiplicity; B--the orthogonal transforming matrix the columns of which are the eigenvectors R_0 which transform the orthonormed system; B^{-1}--the matrix inverse to B; B'--the transposed matrix.

From (2.160) we have

$$R_0 = BL^{1/2} L^{1/2} B^{-1} = A_0 A_0', \qquad (2.162)$$

hence $A_0 = BL^{1/2}$, where $L^{1/2}$--the diagonal matrix from the square roots of the eigen numbers.

The solving of the equation (2.162) of the R_0 matrix also comprises the most essential part of the calculation procedures in factor analysis. The geometrically described transformations are the equivalent of rotating the initial system of coordinates in such a manner that the new base axes coincide with the symmetry axes (the main axes) for the distribution of vector X.

Depending upon the method of determining the communality estimates, a distinction is drawn between two variations [16]: 1) the communaltiy estimates are considered equal to one, and this is the so-called closed factor analysis model; 2) the communality estimates are taken below one, calculating them from empirical data (an open model). Let us examine a closed model as a simpler method which has proven itself in a number of practical problems [17].

After determining the factor loads which correspond to the aggregate of unobservable variables (factors), usually an attempt must be made to interpret them, that is, a certain useful and generally accessible interpretation of the essence of various aspects of a complex phenomenon as reflected by the isolated factors.

Due to the fact that the procedure of obtaining the loads in factor analysis does not lead to a uniform result (with a number of factors greater than one), it is possible to obtain equivalent sets of loads by their orthogonal transformation. Geometrically this will correspond to an additional rotation of the factors in the measurement space.

As the criteria for locating the optimum (in the sense of interpretation) position of the factor axes in space, rather many proposals are known. The varimax criterion has proven effective in a number of actual studies [24, 26] and the sense of this comes down to reducing the factor loads to the simplest type. The simplicity V_j of any factor is determined in the given instance as the variance of the squares for the corresponding factor weights:

$$V_j = \left[n \sum_i (a_{ij}^2)^2 - \left(\sum_i a_{ij}^2\right)^2\right]/n^2.$$

The varimax criterion consists of demanding the maximization of the sums

$$V = \sum_j V_j \to \max.$$

For obtaining unbiased estimates, the values a_{ij} are normed by dividing them into the corresponding communalities h_i^2. The final varimax criterion is determined by the ratio

$$V = \sum_j \left\{\left[n \sum_i (a_{ij}^2/h_i^2)^2 - \left(\sum_i a_{ij}^2/h_i^2\right)^2\right]/n^2\right\} \to \max, \qquad (2.163)$$

while the solution is written in the form $A = A_0 T$, where A_0--the matrix of factor weights obtained by the main component method; T--the orthogonal transforming matrix selected in such a manner that the simplicity V of matrix A is maximal.

In the described method the most labor intensive part is the calculating of the eigenvectors of matrix R_0 using the main component method. At present a number of machine programs are known realizing the presented method and its modifications. One such program (compiled by Ye. Yenchenko) is described in [24]. It has been used in analyzing statistical complexes which describe, for example, such an object at the network of USSR trunk air routes. In working out the development forecast for the demand for air passenger traffic in [33], a set of methods was used including trend extrapolation for short lead times, expert questioning to discover possible shifts in demand, and in addition, factor analysis was used for the network of Soviet trunk air routes.

A fragment of the network was studied consisting of 12 air routes holding the first places among the nation's routes in terms of the passenger traffic volume. As the characteristics they took the following statistical data on: 1) the amount of annual passenger air traffic; 2) the amount of passenger traffic by rail between the end points of the connections; 3) the population dynamics of the cities which were the end points of the connections; 4) the ratio of the amount of per capita national income to the air fare for the route by years. Thus, the examined stochastic complex was represented by an aggregate of values for 48 variables.

The computer calculation following the described method (carried out by V. L. Gorelova) showed that the correlation matrix has a very high "rigidity" estimate (0.797) and the number of essential connections is 63.5 percent with a significance level of $P = 0.01$. The calculating of the factor load matrix with the subsequent rotating of the main factors in accord with the varimax criterion (2.163) made it possible to isolate three main factors with a total contribution of 90.3 percent to the generalized sample variance. The interpretation of the isolated factors led to their following explanation in accord with the distribution of the factor loads.

The first factor (77.2 percent of the generalized sample variants), as in a majority of factor research, has high loads for virtually all indicators, and somewhat greater for the indicators of the first and fourth groups and somewhat less for the indicators of the second and third groups. It can be interpreted as the generalized main factor for air passenger traffic. The second factor (7.8 percent) can be interpreted as the rail factor, since the greatest factor loads in it occurred as an average in the area of the second group of indicators. The third factor (5.3 percent) can be termed the population factor, since the greatest factor loads occurred for it in the area of the third group of indicators.

The results of the experiment showed how effective factor analysis is as a tool in studying multivariate stochastic objects. It not only makes it possible to significantly reduce the dimensionality of the description (from 48 to 3 in the designated example) without essential losses of accuracy but also isolates the main factors which disclose the basic driving forces of the process and these can rather easily be interpreted from the positions of human, logical understanding of its essence.

Having examined the use of factor analysis in the retrospection and diagnosis stages in forecast research, let us take up the possible method of applying it in the stage of the immediate elaboration of the forecast (the prospection stage).

The essence of the method of research forecasting based on factor analysis can be presented in the following manner.

According to (2.155) and (2.156)

$$X = AF + U,$$

where $U = I - H^2$--the matrix of specific factors and errors which is closer to zero the closer the communality matrix H^2 is to one.

In the closed model which we have used, it is assumed that H^2 is a singular matrix and, consequently, it can be considered that

$$X = AF. \tag{2.164}$$

By retrospective analysis of the n-dimensional random vector X the matrix was determined for the factor loads A with a dimension of $n \times n$. In examining this matrix it was discovered that only a small number of factors $k < n$ determines the development essence of a multiparametric stochastic system described by the vector X.

Having isolated the corresponding k from the columns of the A matrix, we obtain a matrix A_1 with a dimension of $n \times k$. The equality will correspond to it

$$X \approx A_1 F. \tag{2.165}$$

Let us determine the matrix A^{-1}turned in relation to A in such a manner that $A^{-1}A = 1$. Let us multiply the lefthand side of (2.164) by A^{-1}:

$$A^{-1}X = F.$$

We obtain a matrix of F values with a dimension of $n \times m$, where m--the number of measurements of the vector X in the retrospective period. From this matrix we isolate the lines corresponding to the previously depicted main factors, and we obtain the submatrix F_1. Each line of this submatrix will determine the development process in time for a certain generalized unobservable characteristic f_j of a complex stochastic process under the condition that all m of the measurements are a time sequence. Proceeding from the basic provision of factor analysis, precisely these k of the characteristics $(k < n)$ determine the process as a whole with sufficient completeness considering all the internal statistical relations.

According to the basic principle of research forecasting we can assume that the statistical structure of the forecasted system is preserved in the segment of the lead time $T(A = \text{const.})$ and the basic development trends for the factors as well.

One of the above-listed methods is used to forecast the development of each of the k factors for the set lead time T. As a result we obtain the values additionally of q measurements of F_1, where $q = T/\Delta t_p$, and Δt_p--the pace of measurements in the retrospective period. Then the F_1 matrix with the new q columns assumes the form F_2 with a dimension of $k \times (m + q)$. Having substituted F_2 in (2.165), we obtain the values of X in the retrospective and future period T:

$$X_{p+T} = A_1 F_2. \tag{2.166}$$

The first m of the matrix columns of (2.166) provide the values for all the x_i indicators for the past m measurements and can be used to test the approximation accuracy. The last q of the columns of (2.166) provide the forecasted values for the parameters of x_i at the various moments of the lead period.

If there is no need to calculate the retrospective values of X, the future values can be calculated using the formula

$$X_T = A_1F_{2q}, \tag{2.167}$$

where F_{2q}--a matrix consisting of the last q of the columns in the F_2 matrix.

In conclusion we must reemphasize that the basic advantage in forecasting the development of factors and not individual variables is not even that this substantially reduces the size of the problem ($k \ll n$), but that in the process of forecasting the factors the problem is automatically solved of synthesizing and correlating the singular forecasts of the individual indicators. This is the basic advantage and perspectiveness of using factor analysis methods in research on complex multiparametric stochastic processes.

2.6. Composite Forecasting Methods[4]

The complexity of large technical systems causes difficulties in forecasting their development using any singular method. For this reason in working out scientific and technical forecasts recently ever-greater use has begun to be made of composite forecasting methods (or systems) which synthesize the algorithms of the singular methods in a certain sequence.

In Soviet and foreign forecasting practices, such systems have been worked as the forecast graph method, PATTERN, PROFILE, the selective method, the double tree method, the CPPO, the weighted estimates method, FORCAST, QUEST, the matrix method, RDE, the functional analysis method and a number of others.

In an analysis of composite forecasting methods the basic thing is to define the composition of the procedures and singular methods comprising the system, their standard sequence and the logical rules for making up the systems. Let us examine the best known composite forecasting methods considering this as well as from the viewpoint of the nature of the input and output information.

The forecast graph method has been worked out by a group of specialists from the Cybernetics Institute of the Ukrainian Academy of Sciences under the leadership of V. M. Glushkov. The given system includes the following procedures and methods: the choice of the forecast object, a study of the background, a classification of events, the formulating of the forecast problem, the working out of the general goal of the forecast, an analysis of the hierarchy, the formulating of events, the adoption of the internal structure, the adoption of an external structure, questioning, mathematical processing, a quantitative evaluation of the structure and varification of the results.

The internal structure of the object is a forecast graph and the external are its elements. The technique for constructing the graph is Delphic, that is, by repeated questioning of the experts the close plans of the graph are made to coincide. The

---

[4] Written together with R. I. Peseleva.

given method makes it possible to obtain a forecast of the ultimate goals, the time and probability of their accomplishment.

The input information obtained from expert i includes: the list of premises for the formulated goal S_j, a time estimate of the goal S_j--t_{ij}, the probability estimate of the goal S_j--p_{ij}, a self-assessment of competence β_{ij}, and the degree of certainty in the forecast γ_{ij}.

The information in the output of the method is as follows: the list of the ultimate goals with their premises, the time for achieving the ultimate goals of the graph t_j, the probability of achieving the ultimate goals of the graph p_j. The calculation is made using the formulas:

$$t_j = \frac{\sum_{i}^{K} t_{ij}\beta_{ij}\gamma_{ij}}{\sum_{i}^{K} \beta_{ij}\gamma_{ij}};$$

$$p_j(t) = \frac{\sum_{i}^{K} \beta_{ij}\gamma_{ij}\Pi p_{ijn}(t - t_{ij})}{\sum_{i}^{K} \beta_{ij}\gamma_{ij}},$$

where j--the ordinal number of the ultimate goal;
i--the number of the expert;
k--the number of experts;
n--the number of premise events.

The method makes it possible to carry out a preforecast orientation, to obtain the internal and external structure of the object and to assess quantitatively the given structure. The criterion for selecting the end goals of the graph is the probability of their accomplishment at the set time.

The *PATTERN* system was worked out in 1964 by the U.S. Honeywell Company as a means for aiding the company's leadership in decision taking on the most important questions of setting the future of U.S. military production. But the principles used in this system make it possible to forecast and analyze an enormous number of data in any area of activity. PATTERN stands for Planning Assistance Through Technical Evaluation of Relevance Numbers. The structure of PATTERN consists of the following elements: the choice of the forecast object, the disclosure of the internal patterns, the writing of a scenario, the elaboration of the forecast problem, the elucidation of the general forecast goal, an analysis of the hierarchy, the formulation of goals, the adoption of the internal structure, the adoption of the external structure, questioning, mathematical data processing, a quantitative evaluation of the structure, verification, the elaboration of the resource allocation algorithm, resource allocation and allocation estimate.

PATTERN makes it possible to disclose the forecast orientation, to work out the internal and external structures of the object, to carry out a quantitative assessment of the object's structure and to work out variations for the resource support of the object's elements. The internal structure is the tree of goals. A system of local criteria operates as the external structure.

The levels of the tree contain the following information: the national goals, the measures, tasks, quotas and principles of the systems, the functional subsystems, the designs of the functional subsystems and scientific-technical problems.

The information obtained on the system's output is the following: the list of ultimate goals, the total weights of the goals which are the indicator of their scientific and technical value. The formula for calculating the total weights K_j is of the following form:

$$K_j = \left[\sum_{i=A}^{D} K_{vij}\right]\left[\sum_{i=E}^{F} K_{vij}K_{sij}\right]\cdot\left[\sum_{i=G}^{H} K_{vij}K_{sij}K_{nij}\right],$$

where i--the number of the level;
j--the ordinal number of the goal;
K_v--coefficient of relative importance;
K_s--"state--time" coefficient;
K_n--coefficient of reciprocal utility.

The indicators of scientific and technical value of the goals in the given system have acquired the name of relative importance coefficients. The concluding stage of the given system--recommendations on resource allocation--assumes rational resource aloocation in accord with the level of the relative importance coefficient.

The *PROFILE* (programmed functional indicators for laboratory evaluation) method was worked out by M. Citron in 1965. It is analogous to PATTERN. The internal structure is a tree of relative importance consisting of four levels: a goal of the conflict type, the form of activity, the problems and the program.

The external structure is formed by a system of criteria: sensitivity, promptness, assistance, internal value, the probability of achieving the goal of the program, scientific-technical progress and resources.

A quantitative assessment of the structure is significantly simplified in comparison with PATTERN:

$$R_j = \sum_{i=1}^{n} q_{ij} r_i,$$

where r_i--criterion assessment;
q_{ij}--assessment of element j in terms of criterion i.

The method provides the following: elucidation of the preforecast orientation, the forecasting of the object's internal and external structure, it makes it possible to provide a quantitative assessment of the structure and to give recommendations on rational resource allocation for the purposes of ensuring the forecasted appearance of the object.

Selective method is designed for the forecasting of R&D [Rus.: NIOKR] programs on the sectorial level. A fundamental distinction of the given system's structure from the ones examined above is that the internal structure consists of two graphs --a tree of goals and a stochastic grid which operate sequentially. The external structure is made up out of elements of the tree of goals consisting of four levels: I--the problems (requirements), II--scientific-technical areas, III--scientific-technical problems, IV--R&D programs.

Coefficients for the relative importance of the R&D programs are the result of processing the tree of goals:

$$R^j_{i=n} = \prod_{j,\ i=1}^{n} \sum_{x=\alpha}^{\nu} q_x s_{jx},$$

where i--the number of the level;
j--the number of the goal;
n--the number of levels;
$\alpha, \ldots, x, \ldots, \nu$--the numbers of the criteria;
q_x--the weight of the criterion;
s_{jx}--the weight of goal j in relation to criterion x.

The method makes it possible to disclose the preforecast orientation, to obtain a forecast of the internal and external structure and its quantitative estimate and provides an opportunity to choose alternative programs by solving the stochastic grid and work out recommendations on resource allocation:

$$S_j = \frac{S_{gen} R^j_{i=n}}{\sum_j R^j_{i=n}}.$$

The "double tree" method was worked out by T. Gordon and M. Raffensberger and is designed for forecasting and setting the priorities of both fundamental research and R&D programs. The internal structure of the object is a theoretical tree consisting of question events; the external structure is an experimental tree. The experimental tree confirms or rejects the theoretical one. The value of a R&D program is determined by the degree of its criticalness combined with the possibility of implementation. Using the theoretical tree, the estimates are determined for the relative importance of its elements and using the experimental tree, the possibility of realizing the same elements.

Preference is given to that research which is not only important for solving a technical problem but is also feasible. The method provides an opportunity to make a preforecast orientation and construct the internal and external structures of the object and also indicates the possible direction of estimating it.

The CPPO System. The Center for Long-Range Research and Economic Estimates of the French Ministry of National Defense has worked out a system for the forecasting of applied research and this has been named after the place of its creation.

The goal of the CPPO system is to provide a forecast for future weapons systems and select the proposed research aimed at providing the required performance of the systems.

As the internal structure of the object in the given system there is a "double tree" (a "utility tree" and an "economic tree") with a common lower level. The utility tree is constructed on the basis of the client's demands. The economic tree considers the interests of the executor. The utility tree consists of the following levels: strategic tasks, tactical tasks (functions), potential requirements, subsystems, and elements of research operations (ERO).

The economic graph has four levels: ultimate goals, tasks, means, ERO, that is, the ERO level is common for both graphs.

By the external structure in the given system one understands the degree to which the elements of the trees match certain criterial features. In a quantitative assessment of the structure, indexes are assigned to the elements. With the presence of several criteria features, combinatorial matrices are used for obtaining the element indices. The element indices are an indicator of their relative importance and to this degree correspond to ranks. The algorithm for the quantitative assessment of the structure using the given system is extremely simplified: the indices of the elements lying on the branches which converge on the lower level are added.

The CPPO system makes it possible to obtain a preforecast orientation, to work out a structure of the forecasted method and evaluate it quantitatively.

The weighted estimates method (WEM [Rus.: MVO]) was worked out for forecasting R&D programs on the sectorial level. According to the given system, the object's internal structure is constructed similar to a tree of goals consisting of five levels: common development goals for scientific and technical progress in the sector, the basic development tasks of science and technology in the sector, the basic areas of scientific and technical research, the basic R&D subject, scientific and technical problems. The external structure is formed by elements of the tree of goals, that is, the elements of the preceding level are the estimate criteria for the elements. Such estimates are called weighted:

$$k_\gamma^{i+1} = \sum_{j=1}^{n} k_j^i a_{j\gamma}^{i+1}; \quad \sum_j k_j^i = 1,0; \quad \sum_\gamma a_{j\gamma}^{i+1} = 1,\ 0,$$

where $a_{j\gamma}^{i+1}$--particular specific weight of element with ordinal number γ on level $i+1$ estimated for element j of level i.

The method provides an opportunity to make a preforecast orientation, to construct a forecast of the object's internal structure, to use it as an internal structure and make a quantitative estimate of the object's structure.

The FORCAST System provides forecasting of scientific and technical development for a period of over 10 years and is widespread in the United States for forecasting the development areas of scientific and technical research in terms of weapons systems. The internal structure is worked out in two directions: the tasks of scientific and technical development and the areas of scientific and technical development. The estimates using the "expenditure--effectiveness" criterion comprise the external structure.

FORCAST provides an opportunity to make a preforecast orientation, to work out the object's structure, to provide a quantitative estimate for the structure, and to prepare recommendations on resource allocation for the purposes of ensuring the forecasted appearance of the object.

The QUEST System (quantitative utility estimate for science and technology) was developed by M. Citron in 1966. The forecasting of the object's structure is carried out on the basis of research on the areas of science and the types of technology and constructing matrices of their reciprocal influence. The structure's quantitative estimate is made by a logical correlating of the influence matrices which describe the area of science by assistance coefficients and the type of technology by contribution coefficients. The QUEST System solves all the most important problems in forecasting scientific and technical programs, from providing a preforecast orientation to elaborating recommendations on resource allocation. The given system includes the following procedures and methods: the choice of the forecasting object, background research, a classification of events, the positing of the forecast problem, the elucidation of factors, factor classification, the constructing of influence matrices, the adoption of the internal structure, the adoption of the external structure, questioning, the mathematical processing of the questionnaires, a quantitative estimate of the structure, verification, formulating of the goal function, the formulating of constraints, the solving of a system of linear equations and estimating resource allocation.

The publications on the QUEST System do not give the algorithms for a quantitative estimate of the structure and there is no description of the criterial function in the problem of resource support for the forecasted appearance of the object, but a class of methods for solving the optimization problem is given and these are linear programming methods.

The matrix method has very many common aspects with the QUEST System. Their functional structures are identical. The elaboration of the object's structure in the given system is based upon a classification of factors which influence the ultimate goal and the grouping of independent homogeneous factors into individual complexes. In the matrix method a quantitative estimate of the structure is made by estimating the influence of the complexes on each other and ultimately on the end goal.

In possessing a vector which describes the degree of importance of the ultimate goals and the matrices for the influence of the various factor complexes on the ultimate goals, we obtain coefficients for the relative importance of the factors in the form of vectors. The components of these vectors are the basis by which the resources are allocated:

$$A = \sum_{j=1}^{n} c^j = \sum_{j} \sum_{i} A_i^j \bar{p}_i^j,$$

where A--general resources;
c^j--resources for complex j;
A_i^j--resources for factor i of complex j;
$\bar{p}_i^j$--vector component of factor i in complex j.

The work under the matrix method ends by the solving of a heuristic problem for resource support of the object's appearance.

The RDE Method (forecasting exploratory research and development) was developed by A. E. Nath of the United States. The contents of the given method are somewhat broader than its name (research and development effectiveness), since as a result of employing the given method one establishes the coefficients of the relative importance of systems and their elements (goals) and the resources which are optimally allocated in terms of the effectiveness criterion. The object's structure is worked out in the form of influence matrices according to the system: research program--research tasks n, research tasks--systems k, systems--goals N, goals--research subjects i.

In using this method in sequence one determines the relative importance coefficients for the systems R_k, for the goals R_n and for the subjects R_i:

$$R_\kappa = \sum_{i,j} D[n] \times D[\kappa];$$

$$R_N = \sum_{i,j} D[n] \times D[N];$$

$$R_i = f(R_\kappa, R_N, t, \alpha, K_i, N_i),$$

where D[n]--the matrix of problems (an element of the matrix is the relative importance of each variation of the problem);

D[k]--the matrix of the system (the element of the matrix is the contribution of the given system to solving the given variation of the problem);

D[n]--the matrix of goals (the matrix element is the contribution of the given goal to solving the given variation of the problem);

t--the envisaged time;

α--the degree of completeness;

K_i--the quantity of systems which are assisted by subject i;

n_i--the number of goals which are assisted by subject i.

The resource allocation problem using the given method is a linear programming problem of the type

$$V_T = \sum_i V_i x_i \rightarrow \max$$

with constraints of the type $\sum_i a_i x_i \leqslant b_i$, where V_T--effectiveness of forecasted research program; x_i--resource for subject i; V_i--effectiveness of subject i;

$$V_i = f\left(\frac{\partial p_i}{\partial t}, R_i, p_i, t, k_i, N_i\right),$$

where $\frac{\partial p_i}{\partial t}$ --the rate of change in the probability of success over the designated interval of time. The method provides an opportunity to construct a forecast of the object in the following sequence: to provide a preforecast orientation, to work out the object's structure, to make a quantitative estimate of it, and provide recommendations on resource allocation for the purposes of ensuring the forecasted appearance of the object.

The functional analysis method is used for forecasting new technology according to the production parameters and the allocating of funds for the disclosed series of scientific research and development. The forecasting of the object's structure is carried out using functional subsystems.

In assessing the structure, as factors influencing the financing level of the scientific research and development, the following have been used: an index of progress J_A (the ratio between the progress of the technological parameter with the given financing level and the progress of the same parameter which could be achieved with unlimited financing over the same period of time); the value index J_W (the result of generalizing the two matrices--consumer and technological value).

For modeling the progress index, an exponential function has been used $J_{Ai}(x_i) = (1 - e^{-x_i/D_i})$ with the constraint $\Sigma x_i = c$, where x_i--the financing level of development i; D_i--the financing level which ensures technological progress at a higher rate.

Here the income function has the form of an additive function

$$P = \sum_i J_{Wi} J_{Ai}.$$

This dependence assumes that the additional income from the new investment of funds increases at a retarded rate. The given model of the allocation problem can be accepted as the basis but its realization in the described form makes a great difficulty due to the disparate understanding of the D_i value. The value of the given method consists in basing the adoption of the income functional dependence upon the volume of allocated funds.

We must also examine the *Dean and Howser method* which is not a system in the full sense of this word. With such a method, in using information on the object's structure and its quantitative estimate as the initial data, it is possible to solve the problem of resource support for the forecasted appearance of the object.

For the forecasting of technical systems the expected expenditures and the expected probability of realizing the systems have been used as the initial data. This method envisages a sequential solution to three problems: 1) with a fixed budget and a given set of systems to allocate the budget among the systems in the best manner; 2) with a fixed budget for a subsystem and the given number of subsystems to allocate the budget between the subsystems; 3) for each subsystem to make an optimum choice of the technical approaches to realizing the subsystem.

The following criteria are employed for solving the set problems: the expected cost of implementing the technical approach; the probability of implementing the system with its fixed cost; the expected value of the set of systems with the fixed cost and one of the conditions: a) the systems do not have any priority; b) certain systems have priority; c) all the systems should obtain financing; the probability of implementing the set of systems with the setting of the overall cost.

The dynamic programming method is employed for solving the problems using the Dean and Howser method. The setting of the problems in the publication dealing with the method is given in a general form without bringing out the criterial functions. The value of the given method consists in the formulating of the accepted criteria and assumptions.

An examination of the composite forecasting methods indicates that their specific purpose consists in the following: to provide a preforecast orientation, to work out the object's internal and external structure, to make a quantitative estimate of the object's structure, and to work out recommendations on resource support for the forecasted appearance of the object.

In the stage of making the preforecast orientation in the designated methods, for the purpose of elucidating the internal patterns, use is made of both the writing of a scenario as well as report (analytical) notations, and for the background study, both an analysis of the statistical reporting (factographic analysis) and document analysis as well as the classification of events.

In the stage of working out and adopting the internal and external structure of the object, most often root trees are constructed and namely trees of goals and stochastic grids or nets. In accord with the object's adopted structure, the elements of this structure are goals, events and subsystems. The formulating of the object's structural elements is preceded by the concretizing of the forecasting problem, that is, by the elaboration of a general goal, and an analysis of the hierarchy which discloses the necessary and sufficient number of corresponding structural levels. By the external structure one understands a system of criteria from which the internal structure is evaluated. In certain methods, for example, FORCAST, the criteria are the elements of the internal structure which are initial for the evaluated elements. A different variation for the elaborating of the object's structure is the construction of influence matrices and this is preceded by a disclosure of the factors which influence the object and their classification.

In the stage of the structure's quantitative estimate, the examined methods assume either the questioning of experts or the work of the experts using the commission method, the mathematical processing of the questionnaires and the opinions of the commission members for the purposes of obtaining a reliable general opinion of the representative expert group and the further use of the obtained resources in the calculations using the algorithm for the quantitative estimate of the forecasted object's structure. The quantitative estimate algorithms are determined both by the internal and by the external structure of the object.

At the end of the stage, the methods make it possible to obtain the following spectrum of quantitative estimates: the expected time of carrying out the program, the probability of carrying it out at the designated time, a coefficient for the relative importance of the program, the ranking of programs and program effectiveness.

The stage of resource support for the forecasted appearance of the object in the composite forecasting systems has a varying depth of elaboration. A number of methods are restricted to a quantitative estimate of the structure (the forecast graph method and the double tree method). In certain systems the research allocation problems are solved by heuristic methods (PATTERN, PROFILE and the selective method).

In a majority of systems, optimization problems for resource allocation are solved (FORCAST, QUEST, RDE, the functional analysis method and the Dean and Howser method).

In conclusion, the following conclusions can be drawn: for composite forecasting for BTS development, with a certain adaptation to the forecasting object, it is possible to employ the forecast graph method, PATTERN, PROFILE, the selective method and the matrix method; a number of methods can be recommended for use under the condition of further working out the methods and their components since in the existing publications they have been presented descriptively without an elucidation of the mathematics (the double tree method, CPPO, FORCAST and the functional analysis method).

CHAPTER 3: CRITERIA FOR ESTIMATING EFFECTIVENESS OF LARGE TECHNICAL SYSTEMS

3.1. Principles for Formulating the Effectiveness Criteria of BTS

Chapter 1 gave the essence of technical and economic analysis for the development of large technical systems and this is the basis for their scientific selection and the disclosure of the relationships between the goals, the means for achieving them and the available resources. Technical and economic analysis of the systems represents an aggregate of various types of research: goal, operational, resource, design, criterial and optimization. The present chapter is devoted to criterial research.

The initial premises for carrying out criterial research are: the data of goal and operational research which determines the choice of the goals and the methods of achieving them, the results of design research which determine the constraint vector for the parameter values of the system and the matrix of design relationships; the data of resource research which defines the vector of resource constraints and the matrix of economic relationships. The composition and structure of the criterial function and the systems of disciplining conditions formed in the process of criterial research are influenced by the hierarchical level and stages of the system's life cycle at which the estimate is made.

For example, the effectiveness of transport aircraft system can be estimated in the air transport system of an economic region, in the system of all air transport, in the unified national transport system and so forth. In addition the composition of the estimate criteria of the aircraft systems at the stage of aircraft designing can have definite differences from the system estimate criteria in the stage of taking it out of operation. The initial premises the result of criterial research are shown in Fig. 3.1. The criterial research is related to the concepts of effect and effectiveness. Let us examine these concepts.

The effect (from the Lat. effectus--execution, action) is the result or consequence of certain causes and actions. Effective means producing an effect, leading to the needed results or actual. Hence effectiveness is the results.

In a mathematical interpretation, effectiveness is a relative or specific amount. Effectiveness is the ratio of the effect (result) to the expenditures on attaining it. This definition also includes the effectiveness indicator for various types of systems

$$E_{yd} = Y/Z, \qquad (3.1)$$

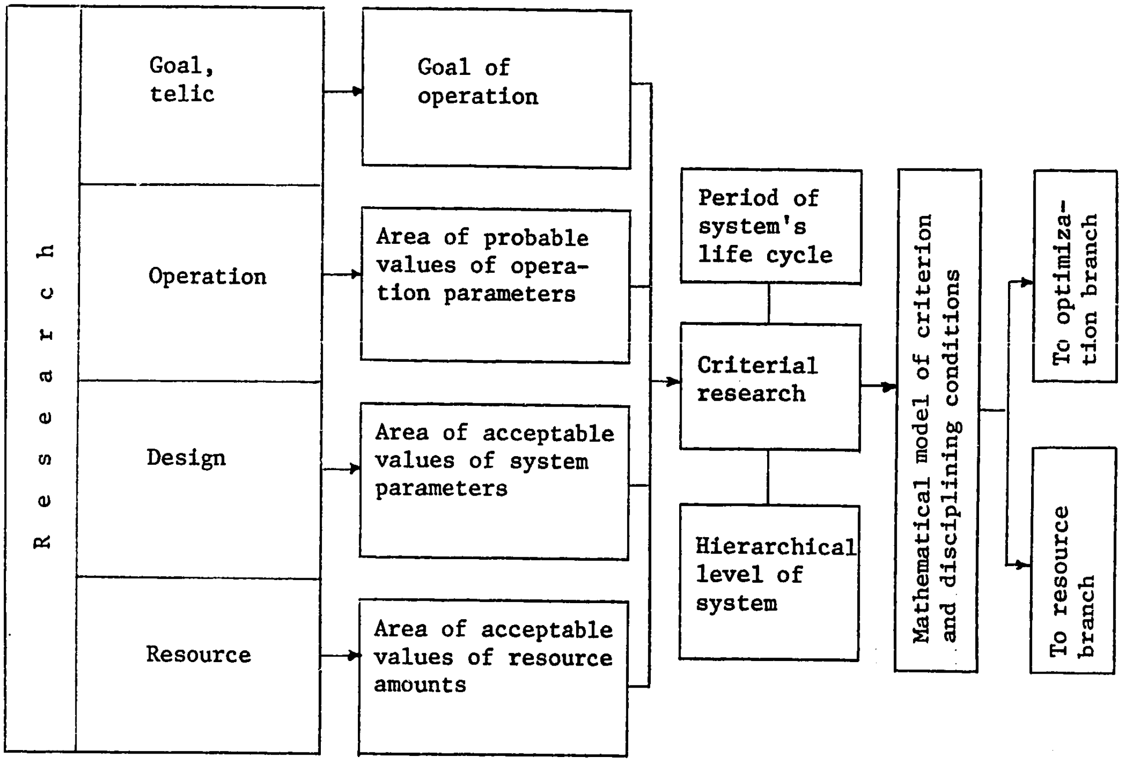


Fig. 3.1. Initial premises and result of criterial research

where Y--the effect (result);
Z--expenditures which provide for the effect Y.

Efficiencies are typical indicators for the effectiveness of equipment.

By effectiveness of large technical systems one must understand the ratio of their effect with the total expenditures in all stages of the system's life cycle (research and development, series production and operation). Thus, the task of studying BTS effectiveness comes down to establishing the amount of the effect, the expenditures and a comparison of them.

The effect of technical systems is sometimes called the technical or target effect. What is the essence of this concept? Any technical system is designed to satisfy a certain social need. A specific, quantitatively determined social need also determines the target effect of the system. For example, any transport system is designed to move freight or passengers. The effect of transport systems can be measured by the quantity of moved freight, by the distance, speed, by traffic safety and so forth.

In order that a system can achieve a certain effect and could satisfy a certain social need, it is essential that it possesses certain properties. For example, aircraft satisfy the needs of the national economy and individuals of our population for air shipments. In order that an aircraft can satisfy such needs, it must reach certain speeds, be light, dependable and so forth. The aggregate of these qualities of an aircraft characterizes its quality.

State Standard 15467-70 defines product quality categories: product quality is the aggregate of product properties determining its ability to satisfy certain needs in accord with its purpose. By analogy with the given definition, it is possible to define the quality of a system as the aggregate of its properties determining its ability to satisfy certain needs in accord with its purpose.

The quantitative characteristics of a system's properties are the indicators of its quality (the parameters). A system's parameters can be divided into three groups: functional, technical and economic. The functional indicators describe the ability of a system to carry out the set functions. The technical and economic parameters describe the structure of the system as well as the consumption of resources involved in its creation, production, storage, transporting and operation.

Functional parameters can be divided into specific-goal and limiting. The specific properties perform a function which determines the possibility of the system to satisfy strictly determined requirements. An example of a specific property of an aircraft is its ability to carry passengers or cargo, that is, its purpose. The most important specific parameter of a system is its productiveness or the property describing the system's ability to carry out a certain amount of useful work over a certain time. Productiveness is a high-rank composite property and it is caused by the comprehensive properties of a lower rank, that is: productivity, reliability, ergonomicness and aestheticness.

Productivity is the ability of a system to carry out a certain amount of work in a unit of time. For example, aircraft productivity describes the maximum transport volume in ton-kilometers carried out in a unit of time (hour, day or year). A system's productivity characterizes its ability to satisfy a certain social need, for example, for passenger or cargo shipments. However, a system can satisfy this need over a brief or extended interval of time and to a varying degree maintain these properties in operation.

Reliability indicators characterize the system's properties to maintain its operating ing efficiency under certain conditions and operating modes. According to the State Standard 13377-75, by the reliability of any object one must understand its property of carrying out set functions in maintaining over time the values of the set operating indicators within the given limits corresponding to the given modes and conditions of use, maintenance, repair, storage and transporting. Reliability is a composite property which, depending upon the purpose of the object and the conditions of its operation, can include the absence of failures, durability, ease of repair and keeping qualities.

By the absence of failures one understands the property of an object to continuously maintain workability over a certain time or certain operating life.

By durability one understands the property of an object to maintain its workability up to a limit state with the established maintenance and repair system.

For example, for an aircraft durability is numerically assessed by the general service life $\tau_{c\Sigma}$. The general service life of an aircraft is formed from the time intervals of its reliable operation or life between repairs τ_{ci}:

$$\tau_{c\Sigma} = \sum_{i=1}^{n_{cp}} \tau_{ci}, \qquad (3.2)$$

where n_{cp}--the number of aircraft major overhauls.

If it is accepted that τ_{ci} = const. = τ_c, then

$$\tau_{c\Sigma} = \tau_c(n_{cp}+1). \qquad (3.3)$$

The repairability of an object is a property consisting in an adaptation for the prevention and detection for the causes of failures[1] and damage and for eliminating their consequences by carrying out repairs and maintenance.

By keeping qualities one usually understands the object's property of continuously maintaining a workable state during and after storage and (or) transporting.

Ergonomicness is a composite property which characterizes the strain and intensity of labor for maintenance personnel. Ergonomicness is determined by a whole series of the system's simple properties, for example, the degree of automation and mechanization of control, the degree of convenient work, the comfortableness of work for maintenance personnel and so forth. Aestheticness is a composite property which influences a person's sensory perception of the entire system as a whole from the viewpoint of its external appearance. A less aesthetic system fatigues a person, it distracts him from the labor process and depresses him. As a result the use of the article deteriorates over time and productiveness drops. Aestheticness is determined by a series of simple properties, for example, shape, design, style and so forth.

A system can maintain its ability to satisfy the set requirements under limited operating conditions. The reflecting of these operating conditions is a function of the article's parameters which reflect the limiting parameters. The limiting properties can be divided into properties which characterize the external conditions of the system's workability and keeping qualities (for example, heat resistance, frost resistance, corrosion resistance and so forth), the properties which characterize production and technical conditions for the system's workability and keeping qualities (for example, maximum speed, maximum altitude and so forth) and the properties characterizing the safety conditions (for example, fire proofness, explosion proofness and so forth).

---

[1] A failure is an event consisting in the violation of a system's workability.

In addition to functional parameters, a system is characterized by technical and economic parameters. In criterial research of the greatest interest are those parameters which in their aggregate determine the consumption of resources related to the development, production, storage, transporting and operation of the system. These indicators can be technical and economic (natural and cost). The technical and economic parameters are manifested in the production sphere, in the distribution sphere and in the consumption sphere.

In production, a system is characterized by one most important property, the design's technological effectiveness. A system can be termed technological in possessing the functional properties which with the set production parameters make it possible to manufacture the system with the minimal expenditures of labor and material resources in the shortest time. In the distribution sphere a system is characterized by those properties such as transportability and storability. The consumability of resources in the process of operating the system is characterized by the resource consumption rate which, in turn, is determined by such simple properties as the labor intensiveness of the performed job, material intensiveness, energy intensiveness and fuel intensiveness. It is also essential to isolate such system properties as operational and repair efficiency.

The above-examined three groups of system quality parameters (functional, technical and economic) in quantitative terms express the corresponding system properties. The functional indicators of the system's quality determine the sphere and conditions of its use and also determine the degree of interchangeability with other systems. As for the economic parameters, they determine the system's quality level and the level of its effectiveness. However, as was pointed out above, the system's quality level is characterized by a large number of parameters and for this reason the system's effectiveness estimate on a basis of them is an extremely complicated problem.

One of the basic purposes of criterial research is to disclose the basic system quality indicators and establish the most important of them. These main indicators are transferred to the category of criteria which must serve as the measure for comparison, estimate and final choice of the system in the stage of optimization research. The remaining indicators can act in the role of disciplining conditions. Obviously the choice of the system's estimate criteria and the setting of disciplining conditions are not a formal but rather a creative procedure which should be carried out using logical analysis, intuition and experience of the leader.

At present, great attention, both in Soviet and foreign literature, is given to the questions of selecting the estimate criteria for technical systems. One of the most debated questions here is the number of criteria required to solve the problem of optimizing the appearance of the BTS.

Until recently, the concept of the effectiveness of technology and technical systems was essentially identified with the concept of economic effectiveness. This is confirmed by the names of the official documents published on the given question, for example, the 1963 "Procedure for Determining Economic Effectiveness from the Introduction of New Technology, Mechanization and Automation of Production Processes in Industry," the 1964 "Basic Procedural Provisions on Determining the Economic Effectiveness of Scientific Research," the 1969 "Standard Procedure for Determining the

Economic Effectiveness of Capital Investments" and others. There is extensive literature on the question of the economic effectiveness of capital investments, new technology, scientific research and so forth. However in recent years in a number of works, along with the economic effect, other effects (results) have been examined for scientific research, development and new technology, including: scientific-technical, social and so forth. Such a positing of the question is correct qad up-to-date.

While previously, in developing machines and the implements of labor, in taking decisions about their development and introduction it was possible to limit oneself to the indicator of their economic effectiveness, but this is not sufficient under present-day conditions. Modern, large technical systems are characterized by extreme complexity of design, of the processes of development, manufacturing and operation, as well as by the importance and complexity of the tasks performed by them.

For this reason, in taking decisions about the development and introduction of such systems, along with the economic effectiveness indicator, it is essential to consider the social consequences (the providing of sanitary-technical and aesthetic production and operating conditions, safety measures and environmental conservation), its scientific-technical level and its influence on the scientific-technical potential of a nation as well as the time needed to carry out the scientific research and prototype design work, the putting of the system into production and operation, the availability of production and human resources, the raw material supply and so forth.

Consideration of the diverse factors indicates that the quantitative basis for decision taking is a complicated problem and the problem of choosing the criterion on the basis of which the BTS effectiveness is determined is also complicated. Depending upon the number of criteria it is possible to speak about two ways of posing the problem of choosing the system's parameters or evaluating it: mono- and polycriterial.

With the monocriterial approach, the evaluation and choice of the technology are made for just one criterion. However, in a majority of decision selection problems there are several criteria which must be considered for correctly selecting the optimum BTS. This is particularly characteristic for the problems of optimizing large technical systems, where a triad of criteria is widely used: effectiveness--cost--time (W--C--T). In a majority of instances, the concent of "effectiveness" is identified with the system's specific effectiveness.

In the case of a monocriterial approach, one of the triad's criteria operates as the main one and the other two as constraints. The question of selecting the main criterion is settled individually in each case depending upon the positing of the optimization problem. For example, if the system's effectiveness has been selected as the main criterion, the problem of optimizing the BTS arrangement is a problem of resource optimum allocation and has the form

$$\begin{cases} W \to \max; \\ C \leqslant \bar{C}; \\ T \leqslant \bar{T}. \end{cases} \tag{3.4}$$

In the case of selecting the system's cost as the main criterion, the system optimization problem comes down to the problem of resource minimization and this is written in the form

$$\begin{cases} C \to \min; \\ W \geqslant \overline{W}; \\ T < \overline{T}. \end{cases} \tag{3.5}$$

In the expressions (3.4) and (3.5) $\overline{W}$, $\overline{C}$, $\overline{T}$--the set limit values of W, C and T.

The mathematical model of the optimization problem, when time is taken as the main criterion, has the form:

$$\begin{cases} T \to \min; \\ W > \overline{W}; \\ C \leqslant \overline{C}. \end{cases} \tag{3.6}$$

In the given general arguments a number of additional disciplining conditions have not been pointed out and these are determined by the essence of the problem being solved and are brought out in the stages of the goal, operational, design and resource research.

Many researchers, in being aware of the need to consider several criteria in optimizing the configuration of technical systems, have followed the path of designing component criteria in the form of various functions from the initial criteria. The use of component criteria within a monocriterial approach to the optimization problem has made it possible (or more often, the illusion of such a possibility) for indirectly considering several optimality criteria.

The component "fractional" criterion has become rather widespread in economic practive and in the problems of optimizing the parameters of new equipment. This criterion is the ratio of the obtained effect to expenditures (3.1) related to the obtaining of the given effect or the inverse ratio.

A fractional criterion is employed in a monocriterial positing of the optimization problem for the purpose of optimizing the system immediately for two criteria: specific [telic] and economic effectiveness.

The use of a fractional criterion has two most essential shortcomings which are related to the form of its presentation. In the first place, in maximizing the proportional effect or minimizing proportional expenditures, no consideration is given to the amount (physical or economic) of the numerator and demoninator. The second drawback of a fractional criterion is that the size of the ratio moves rapidly toward its limit values.

In summarizing what has been said above, it can be concluded that the basic shortcoming of the monocriterial approach to solving optimization problems in constructing technical systems consists in the complexity and subjectivism of choosing the main optimality criterion. Among the merits of the monocriterial approach is the possibility of extensively utilizing the rich arsenal of mathematical methods from decition taking theory, namely: mathematical programming, game theory, the theory of statistical decisions and various numerical optimization methods. The polycriterial approach is essential in the following types of decision choosing problems: 1) when the choice of the optimum variation of a system involves considering the effectiveness of the component subsystems, each of which is estimated by an individual criterion; 2) when the quality of the decision must be assessed for a number of variations of conditions for the use of the technical system and an individual estimate is introduced for each variation; 3) when the decision is evaluated over time or by stages in the BTS life cycle and an independent criterion is introduced at each stage; 4) when there are several goals in the functioning of the BTS.

An example of considering several effectiveness criteria is the forms of additive (3.7) and multiplicative (3.8) criteria which have been widely used in optimization practices:

$$E_1 = \sum_{i=1}^{m} \lambda_i e_i; \qquad (3.7)$$

$$E_2 = \prod_{i=1}^{m} e_i^{\beta_i}, \qquad (3.8)$$

where E_1 and E_2-- component criteria;
e_i--local i criterion;
λ_i, β_i--significance and elasticity coefficients for local i criterion in the composite criterion;
i--ordinal number of criterion i $\in$ [1; m].

These types of criteria, particularly the additive one, have been recommended in the literature and are used in practice, however the subjectivity of assigning the λ_i and β_i coefficients (particularly in the case of the local criteria for differing physical dimensionality) reduces the value of the recommendations obtained in accord with the criteria (3.7) and (3.8) for choosing optimum decisions.

Recently the problem of polycriterial (another widespread name is vector) optimization has assumed timely and exceptionally important significance. The problem of vector optimization in a determined case (that is, in the absence of random and undefined factors) in a general form can be formulated in the following manner.

Let there be a certain operation, the outcome of which is estimated by the aggregate of local criteria e_1, e_2, ..., e_m forming the vector of the effectiveness criteria $E = f(e_i)$, $i \in [1; m]$. The local criteria of e_i can be both scalers and vectors. The relative importance of the local criteria is set in the form of the vector $\Lambda = (\lambda_1, \lambda_2, ..., \lambda_m)$ the specific sense and importance of which are not important in the general positing of the problem.

The outcome of the operation depends upon the values of the strategy (decision) for the operating side. Strategy X can be a scalar, vector or a matrix. The operating side controls the operation, choosing stragegy X from the Ω_X of its tolerable values. The area of Ω_X is set by a certain aggregate of disciplining conditions.

The vector of the E criteria is related to solving the X of the mapping F: $X \rightarrow E = F(X)$, where $F = (f_1, f_2, \ldots, f_m)$. The mapping can be set analytically or in a more complicated instance algorithmically.

In the problem it is essential to find the optimum value of X^o which is determinable by two conditions: 1) the solution should be feasible, that is, it should belong to the set Ω_x of its tolerable values; 2) the solution should be the best, that is, to optimize the vector of the effectiveness criteria E considering the relative importance vector Λ.

In other words, an optimum solution should satisfy the ratio

$$X^o = F^{-1}\left[\operatorname*{opt}_{X\in\Omega_x}(E(X),\ \Lambda)\right], \qquad (3.9)$$

where F^{-1}--the inverse mapping $E \rightarrow X = F^{-1}(E)$;

opt--a certain optimization operator adopted in the problem.

Let us give an example of a task of decision taking with a vector effectiveness criterion. It is essential to choose a model of a multipurpose aircraft. This is to perform simultaneously the functions of a transport (misfunction is estimated by the e_1 criterion), passenger (the e_2 criterion) and ambulance aircraft (e_3) and so forth. The quality of the multipurpose aircraft can be estimated only by the vector criterion $E = (e_1, e_2, \ldots, e_m)$.

In solving vector problems of decision taking a number of specific problems arise having not a formal or computational nature but rather a conceptual one. The main one is selecting the optimality principle which determines the optimum decision properties and provides an answer to the main question of why the optimum decision is the best of all the other decisions (surpasses the other decisions). In the model of the problem (3.9) this corresponds to disclosing the sense of the operator opt E.

In the problems of monocriterial scalar optimization, the optimality principle is uniform for all the problems. As the optimum solution X^o a value of X is taken for which the condition is valid (in the case of the maximization of the criterion)

$$e(X^o) \geqslant e(X) \text{ for all } X \in \Omega_x, \qquad (3.10)$$

where e--scalar effectiveness criterion.

The major distinction of vector optimization problems is that for them there is a multiplicity of different optimality principles leading to the choice of different optimum solutions. This places serious demands on selecting the optimality principle. Let us list the basic problems related to solving a vector optimization problem.

Problem 1. A determining of the area of compromises or decisions optimal according to Pareto. In vector optimization problems there is a contradiction between certain of the criteria. This contradiction obviously is not strict as otherwise the problem would not have a solution. Because of this the area Ω_x of acceptable solutions is broken into two nonintersecting parts: the area of agreement Ω_x^c and the area of compromises Ω_x^k. In the area of agreement Ω_x^c there are no contradictions between the criteria and the quality of the solution can be improved simultaneously for all criteria. In the compromise area Ω_x^k there is a contradiction between certain criteria: an improvement in the quality of the solution for some criteria worsens the quality of the solution for others.

Obviously, the optimum solution can lie only in the compromise area, that is, $X^o \in \Omega_x^k$, as in the agreement area the solution can be improved for all indicators. Consequently, the search for the optimum solution must be restricted solely to the compromise area Ω_x^k. Hence Problem 1 is the isolating of the compromise area Ω_x^k from the area of acceptable solutions of Ω_x. In individual instances the search for the optimum solutions with a practically acceptable accuracy can be restricted to isolating the compromise area without a further solution, that is, without the solution to problems 2 and 3.

Problem 2. The choice of the optimality principle and the corresponding compromise scheme. The further search for optimum solutions in the compromise area can be carried out only on the basis of a certain compromise scheme. The number of possible compromise schemes is very great. The choice of the compromise scheme is a complicated conceptual problem.

The choice of the compromise scheme corresponds to disclosing the sense of the optimization operator opt in (3.9) usually in the form

$$\underset{x \in \Omega_x}{\mathrm{opt}}\, E(X) = \underset{x \in \Omega_x^K}{\mathrm{opt}}\, E(X) = \underset{x \in \Omega_x^K}{\max}\, \varphi(E(X)), \tag{3.11}$$

where $\phi(E)$--a certain scalar function of the criteria.

Problem 3. Normalization of the criteria. In practical terms this problem comes down to adjusting the chosen compromise scheme.

The listed problems are the main ones but they do not exhaust the entire range of problems. These problems (with the exception of the first) are of a conceptual nature. The solution to similar problems should be carried out using strictly elaborated formalized procedures on the basis of scientific arguments, with the limited and formalized use of heuristic procedures.

Let us take up in somewhat greater detail the central problem of vector optimization, that is, the choice of the compromise scheme. At present there are no sufficiently strict theoretical concepts on the choice of optimum compromise solutions, that is, for determining the scheme to locate the optimum variation of the system in the compromise area. Certain authors have proposed restricting oneself to just isolating the compromise area while the decision is made proceeding from the subjective views of individual responsible persons. With such an approach the subjective factor has the dominating influence in decision taking and this leads to its insufficient soundness. Moreover a whole series of schemes have been proposed for compromise solutions where the crucial role has been assigned to experts in carrying them out. Such a positing of the question is more valid. Scientifically sound expert evaluation possesses sufficient objectivity in decision taking.

Let us take up certain schemes for locating compromise solutions. The evenness principle consists in an even rise simultaneously in the level of all the local criteria. The conditions for its use are: the local criteria are normed (that is, they have a uniform scale of measurement) and equal in terms of importance. There are several varieties of this principle, for example, the equality and maximin principles. In accord with the equality principle, maximization is achieved under the condition of the equality in the level of all criteria. The solution with such an approach can be beyond the compromise area. The other variety of the evenness principle, the maximin principle, in individual instances is employed in determining perspective systems. Here the goal is a desire to obtain a satisfactory result for all criteria by bringing up the poorer of the criteria.

To a greater degree the principle of a "just" concession fits a solution to the problems involved in determining perspective BTS and it is based upon an assessment and comparison of the increase and loss of local criteria inevitable in the compromise area. This principle has two varieties: the principle of an absolute concession and the principle of a relative concession.

The absolute concession principle states: a compromise is valid when the total absolute level of the decline in one or several criteria does not exceed the total absolute level in the rise of the other criteria. This principle is met by a model for the maximization of the total criteria (integral effectiveness)

$$\operatorname{opt} E = \max \sum_{i=1}^{m} e_i. \qquad (3.12)$$

The shortcoming of the absolute concession principle is that it can permit a sharp differentiation in the levels of the individual criteria since a great value of the integral criterion can be achieved by a high level of some criteria with a low level of the remainder.

The relative concession principle states: a valid compromise is the one where the total relative level in the decline of one or several criteria does not exceed the total relative level for the increase in quality for the other criteria. The relative concession principle is met by the optimization model with a criterion in the form of the product of the local criteria

$$\operatorname{opt} E = \max \prod_{i=1}^{m} e_i. \qquad (3.13)$$

The relative concession principle is very sensitive to the amount of the criteria as due to the relativity of the concession there is an automatic decline in the price of the concession for the high-value criteria and vice versa. As a result there is a significant smoothing of the local criteria levels.

There also is a whole series of other optimality principles. We feel that at present, for solving problems related to choosing the optimum BTS, the principle of isolating the main criterion is the soundest scientifically and the most elaborated. Here definite constraints should be imposed on the established range of local criteria. Such an optimality principle can be termed a polycriterial positing of the problem with constraints. In individual instances, in realizing the given principle, it is advisable to utilize the equivalent reciprocal models in which the main and most important local criteria which are in the constraint system can change places.

Optimization is carried out on a basis of the maximization (or minimization) of the local criterion chosen as the main one and a certain aggregate of local criteria which reflect the most important functional, economic and time characteristics involved in the creation, production and operation of the BTS.

In estimating the alternate BTS, one of the triad of criteria of (3.4)-(3.6) can be adopted as the global criterion. The most typical is the forming of a resource minimization criterion. In this case as the criterial function one adopts the economic effectiveness criterion with a constraint on the telic effectiveness and time. If it is considered that there presently are economic models which take the time factor into account as well as methods for a value assessment of time, considering this the above-given idea can be formulated in the following manner. In resource minimization, as the criterial function, one adopts the economic effectiveness criteria with a constraint on the telic effectiveness and considering the time factor. Let us examine in greater detail the essence of the economic effectiveness criterion for large technical systems.

3.2. Economic Effectiveness Criteria and Types of Economic Effects of BTS

For creating BTS, great capital investments are required in all stages of its life cycle. The program for the development of BTS to a certain degree is a long-term capital investment program or a long-term investment program and for this reason the economic effectiveness of the BTS can be viewed as the effectiveness of capital investments into these systems. Here it is possible to utilize the entire methodological apparatus for assessing the economic effectiveness of catpial investments.

The methods for calculating capital investment economic effectiveness are given in the "Standard Procedure for Determining the Economic Effectiveness of Capital Investments" as approved by the Decree of the USSR Gosplan, the USSR Gosstroy and the Presidium of the USSR Academy of Sciences of 8 September 1969. Capital investment effectiveness is determined by a comparison of the effect and the expenditures. In planning and designing, the overall (absolute) economic effectiveness is determined as the ratio of the effect to the total capital investments and in selecting the models for solving economic or technical problems this is the comparative economic effectiveness showing to what degree one model is more effective than another.

In the Standard Procedure the effectiveness criteria (indicators) are examined for the system levels:

a) For the national economy as a whole, for the Union republic economy and the national economic sectors (industry, agriculture, transportation and construction) as the ratio of the increment in the volume of national income (net product) with its set physical structure in comparable prices ΔD to the capital investments K into the material production sphere which have caused this increase:

$$Y_{ci} = \Delta D:K;$$

b) For the individual sectors and subsectors of industry, agriculture, transportation and for the ministries, by the ratio of the increase in profit to the capital investments causing this increase:

$$Y_{ci} = \Delta R:K,$$

where ΔR--the increase in the annual profit over the planned period (year, 5-year period);
K--capital investments into the construction of production-end projects;

c) For individual enterprises, construction projects and sites, for individual measures and technical-economic problems, by the ratio of profit to the capital investments:

$$Y_c = (P - C):K, \tag{3.14}$$

where P--the value of the annual product output (according to the plans) in enterprise wholesale prices (without the turnover tax);
C--the costs of annual product output;
K--the estimated cost of the project under construction (capital expenditures on carrying out the measures and technical-economic problems).

In calculating overall economic effectiveness the repayment time is determined for the total volumes of capital investments on the basis of the inverse ratio of capital investments or profit or the savings from the reduction in production costs:

$$T_{ci} = K:\Delta R; \quad T_c = K:(P - C); \quad T_{cc} = K:(C_1 - C_2).$$

Calculations for comparative economic effectiveness of capital investments are employed in comparing the variations of economic or technical decisions, in locating enterprises and their complexes, in solving problems related to choosing interchangeable products, for the introduction of new types of technology, for the construction of new enterprises or the reconstruction of existing ones and so forth.

The indicator for the comparative economic effectiveness of capital investments is a minimum of reduced expenditures. The reduced expenditures for each variation are the total current expenditures (production costs) and capital investments reduced to a uniform dimensionality in accord with the normed effectiveness coefficient:

$$C_i + E_n K_i = \min, \tag{3.15}$$

where C_i--current expenditures (production costs) for each variation;
E_n--the normed capital investment effectiveness coefficient;
K_i--capital investments for the same variation.

According to the Standard Procedure for the national economy a normed capital investment effectiveness coefficient has been set equal to 0.12 and this corresponds to a normed repayment time $T_n = 8.3$ years, where $T_n = 1/E_n$.

In February 1977, the State Committee of the USSR Council of Ministers for Science and Technology, the USSR Gosplan, the USSR Academy of Sciences and the State Committee of the USSR Council of Ministers on Inventions and Discoveries approved the "Procedure (Basic Provisions) for Determining the Economic Effectiveness of Using New Technology, Inventions and Rationalization Proposals in the National Economy." According to this procedure a uniform normed coefficient for the economic effectiveness of capital investments was set at 0.15, and this corresponds to a normed repayment time $T_n = 6.7$ years.

The normed capital investment effectiveness coefficient is the lower limit of capital investment effectiveness. The amount E_nK is the normed savings. This is the amount of the national economic losses which arise not at the enterprises which carry out the capital investments but at other enterprises which have not received the necessary capital investments since they were used in the given element of the national economy. If one examines a certain technical system, the lost net income in the other sectors for the designated system is additional expenditures caused by the investments into it.

The indicator of reduced expenditures is identical to the indicator for the repayment time of additional capital investments:

$$T_{2/1} = \frac{K_2 - K_1}{C_1 - C_2},$$

where $T_{2/1}$--the repayment time of the additional capital investments into the second variation in comparison with the first due to the savings in operating expenses, years;

K_1, K_2--capital investments in the first and second variations of the design, rubles;

C_1, C_2--production costs (operating expenditures) in the first and second variations of the design, rubles/year.

The second variation will be more economic than the first under the condition $T_{2/1} \leqslant T_n$, where T_n--the normed repayment time of the additional capital investments, years.

Let us illustrate the principle of selecting the most effective variations in the given example. Let us assume that there are two variations for the construction of a plant and the indicators for these are shown in Table 3.1.

Table 3.1

Name	Indicators, million rubles, by variations	
	1	2
Volume of annual product	$P_1 = 24$	$P_2 = 24$
Capital expenditures	$K_1 = 15$	$K_2 = 30$
Product costs	$C_1 = 19$	$C_2 = 15$
Profit	$P_1 - C_1 = 5$	$P_2 - C_2 = 9$
Reduced expenditures	$C_1 + E_nK_1 = 21.3$	$C_2 + E_nK_2 = 19.5$
Indicator of total effectiveness of capital expenditures	$Y_{c1} = \frac{P - C_1}{K_1} = 0.33$	$Y_{c2} = \frac{P - C_2}{K_2} = 0.30$

Let us select the most effective variation. With the same volume of sold product, the second variation requires 15 million rubles more of capital investments (K_1-K_2). However the second variation of the plant provides a reduction in production costs in comparison with the first variation by 4 million rubles (C_1-C_2). Let us calculate the repayment time for the additional capital investments:

$T_{2/1} = (K_2-K_1)/(C_1-C_2) = 15:4 = 3.75$ years. Thus, $T_{2/1} < T_n (3.75 < 6.7)$, so we can conclude that the second variation of the plant is better than the first.

The given method for selecting the variations according to the formula for the repayment time for the additional capital investments can be employed with a small number of variations. The basic difficulty of this method is that the variations must be compared in pairs: the first with the second, the third with the second and so forth. Because of this the method of reduced expenditures is widely used [formula (3.15)]. In the given example the second variation is also chosen for the minimum of reduced expenditures (19.5 million rubles).

Above we have examined the indicators of general and comparative capital investment effectiveness as given in the Standard Procedure. The Standard Procedure points out that the designated uses of capital investments can be considered economically effective under the condition that the obtained general effectiveness coefficients are not below the planned norms and the analogous indicators for the preceding plan period. In the given example (see Table 3.1), both variations are effective, since the general effectiveness coefficient (0.33 and 0.30) exceed the norm for capital investment effectiveness in industry which we will consider as equal to 0.15. The question arises why are the comparative effectiveness indicators required? Is it not possible to select the most effective variation using the maximum general effectiveness indicator? If the variation is chosen using this criterion, then preference must be given to the first variation (0.33 0.30). A contradiction develops as according to the general effectiveness criterion the most effective is the first variation while according to the reduced expenditure indicator it is the second. Which solution is correct?

Let us attempt to select the most effective variation using the maximum general effectiveness indicator. The maximum value of this indicator (0.33) corresponds to the first variation. Let us compare both variations. The second variation makes it possible to obtain 4 million rubles of additional profit per year in comparison with the first variation and here the general effectiveness indicator is 0.30, that is, only 0.03 less than the analogous indicator for the first variation. Previously it was stipulated that the actual effectiveness of analogous plants for all industry is 0.15. The question arises of why, if there are no capital investment constraints, is it necessary to give up 4 million rubles of profit a year with a general effectiveness of 0.30 which is significantly more than the normed amount (0.15).

The second variation makes it possible to obtain an addition 4 million rubles of profit per year in comparison with the first but requires 15 million rubles more of capital investments. The effectiveness of the additional investments is 4:15 = 0.26, and this is significantly above the adopted normed coefficient for capital investment effectiveness ($E_n = 0.15$). Hence, one must give up the 4 million rubles of additional profit, as the effectiveness of the additional capital investments is below the capital investment effectiveness for the first variation for the basic capital investments.

The general effectiveness indicator for the second variation can be represented in the following form:

$$\frac{P-C_2}{K_2}=\frac{P-C_2+C_1-C_1}{K_2}=\frac{\dfrac{(P-C_1)K_1}{K_1}+\dfrac{(C_1-C_2)(K_2-K_1)}{K_2-K_1}}{K_2}$$

or

$$\frac{P-C_2}{K_2}=\frac{Y_{c1}K_1+E_{2/1}(K_2-K_1)}{K_2},$$

where $E_{2/1}$--comparative effectiveness coefficient for second variation in relation to the first.

In the designated example

$$\frac{P-C_2}{K_2}=\frac{0.33\cdot 15+\dfrac{(19-15)(30-15)}{(30-15)}}{30}=0.30.$$

Thus, in comparing the two capital investment variations, the general effectiveness indicator for one of them is the average weighted (for the capital investments) amount of the general effectiveness indicator of the other variation and of the comparative effectiveness of the first variation in relation to the second. Since $E_{2/1}>E_n$, the additional capital investments under the second variation are effective and, consequently, so is the second variation as a whole. The choice of the second variation is confirmed by the minimalness of the reduced expenditures (19.5 million rubles). In adopting the second variation we maintain the effectiveness of the first and, in addition, we utilize the additional capital investments with an above-norm effectiveness. Consequently, in choosing the capital investment variation it is essential to use the comparative effectiveness method.

There is also another viewpoint on the given question. Thus, individual economists recommend choosing the optimum variation from the maximum indicator of general effectiveness. From the variations given in Table 3.1, the representative of this approach would choose the first variation, arguing as follows. For the 30 million rubles (capital investments in the second variation) when necessary one could build not one but two plants under the first variation, double the product volume and obtain an additional 1 million rubles of profit with the same capital investments. Table 3.2 gives these variations.

Table 3.2

Name	Indicators, million rubles, by variations	
	1	2
Volume of annual product	$P_1 = 48$	$P_2 = 24$
Capital expenditures	$K_1 = 30$	$K_2 = 30$
Product costs	$C_1 = 38$	$C_2 = 15$
Profit	$P_1 - C_1 = 10$	$P_2 - C_2 = 9$
Reduced expenditures	$C_1 + E_n K_1 = 42.5$	$C_2 + E_n K_2 = 19.5$
Indicator of general capital investment effectiveness	$Y_{c1} = \frac{P_1 - C_1}{K_1} = 0.33$	$Y_{c2} = \frac{P_2 - C_2}{K_2} = 0.30$

At first glance the above-given arguments are convincing. The first variation makes it possible with the same capital investment volume to double the product volume and obtain 1 million rubles of additional profit and a higher general effectiveness. However, such a comparison is not objective as it puts the second variation under intentionally poorer conditions in terms of the production volume. Let us eliminate this shortcoming. In actuality, if the demand for the product is 48 million rubles, then for producing it it would be possible to build not only two plants for the first variation but also two plants for the second (Table 3.3).

Table 3.3

Name	Indicators, million rubles, by variations	
	1	2
Volume of annual product	$P_1 = 48$	$P_2 = 48$
Capital expenditures	$K_1 = 30$	$K_2 = 60$
Product costs	$C_1 = 38$	$C_2 = 30$
Profit	$P_1 - C_1 = 10$	$P_2 - C_2 = 18$
Reduced expenditures	$C_1 + E_n K_1 = 42.5$	$C_2 + E_n K_2 = 39.0$
Indicator of general capital investment effectiveness	$Y_{c1} = \frac{P - C_1}{K_1} = 0.33$	$Y_{c2} = \frac{P - C_2}{K_2} = 0.30$

In the designated instance, preference must be given to the second variation which is characterized by minimum reduced expenditures (37.2 million rubles). This variation makes it possible to obtain 8 million rubles of profit in comparison with the first variation with a coefficient of $Y_c = 0.3$ which is only 0.03 less than the corresponding indicator for the first variation. From the additional profit, the additional capital investments under the second variation (30 million rubles) are repaid in 30:8 = 3.75 years and this is significantly less than the accepted norm (6.7 years).

Thus, the most effective variation is selected from the comparative effectiveness indicator. Here it is assumed that the capital investment volume can be increased to 60 million rubles. However, if the capital investment volume is restricted to 30 million rubles, the persons making the decision should choose the first, less effective variation. Thus, the volume of allocated capital investments operates as a constraint in choosing the variations of the system.

It would be wrong to limit oneself to just the comparative effectiveness indicators. The comparative effectiveness indicator makes it possible to determine that one variation is more effective than another but it does not assess the general effectiveness of each of them. Let us explain this by the example given in Table 3.4.

Table 3.4

Name	Indicators, million rubles, by variations	
	1	2
Volume of annual product	$P_1 = 27$	$P_2 = 27$
Capital expenditures	$K_1 = 80$	$K_2 = 110$
Product costs	$C_1 = 19$	$C_2 = 15$
Profit	$P_1-C_1 = 8$	$P_2-C_2 = 12$
Reduced expenditures	$C_1+E_nK_1 = 31.0$	$C_2+E_nK_2 = 31.5$
Indicator of general capital investment effectiveness	$Y_{c1} = \frac{P-C_1}{K_1} = 0.10$	$Y_{c2} = \frac{P-C_2}{K_2} = 0.11$

In the example the first variation is the most effective as this provides a minimum of reduced expenditures (31.0 million rubles). The additional capital investments (K_2-K_1) = 30 million rubles are repaid from the reduction in production costs (C_1-C_2) = 4 million rubles in 7.5 years and this is above the normed time (6.7 years). However neither the first nor the second variation is absolutely ineffective and this is confirmed by the values of the Y_{c1} and Y_{c2} coefficients, each of which is less than the adopted normed coefficient for capital investment effectiveness (0.15). If we view comparative effectiveness as the repayment rate of additional capital investments and general effectiveness as the repayment of full capital investments, then in the designated example neither in the first nor in the second variation are the full capital investments repaid at the normed time and consequently both variations are economically ineffective. If the construction of such a project is essential to the national economy and the economy decides to build a knowingly unprofitable project, then the most effective variation of the two ineffective ones is chosen using the comparative effectiveness indicator which in the given instance demands a greater effectiveness level from the additional capital investments than the full capital investments.

From the above-given arguments it can be concluded that the standard procedure correctly recommends the use of both the general effectiveness indicator and the comparative effectiveness indicator in the design stage. The determining of these indicators can be combined in the following sequence. From all the capital investment variations one choses those for which the general effectiveness indicators are above the normed amount. Out of all the absolutely effective variations, the most effective is the one for which the amount of reduced expenditures is minimal.

The proposed indicators are employed in observing definite conditions. The compared variations should be compatible in terms of product quality and quantity, in terms of the prices employed for expressing the expenditures and the effect and in terms of the time of making the expenditures and obtaining the effect. Of course, the reducing of the variations to a comparable type for the listed features is sometimes a complicated task. However this reduction must be carried out in order to obtain an objective estimate of each variation. Let us take up in more detail the comparability of variations for expenditures made at different times.

Expenditures which are the same in terms of amount but made at different times are economically not equivalent. Let us assume that there are two variations for investing the same amount of expenditures. According to the first variation the project is to be built in 5 years and according to the second in 3 years. According to the first variation the capital investments do not produce any return for 5 years and for the second for 3 years. If the second variation is adopted, then for 2 years a certain amount of capital investment is freed which could be channeled into other projects and from this the national economic outlays could be additionally reduced. In turn, the obtained savings in being productively employed also will make it possible to reduce national economic expenditures and so forth.

If at the initial moment of time the amount of capital investments to be additionally put into economic circulation is K_o, then the effectiveness of its use will be:

at the end of the first year of the planned period

$$K_1 = K_o + K_oE_n = K_o(1 + E_n);$$

at the end of the second year of the planned period

$$K_2 = K_1 + K_1E_n = K\ (1 + E_n)(1 + E_n) = K_o(1 + E_n)^2;$$

at the end of year t of the planned period

$$K_t = K_o(1 + E_n)^t.$$

The multiplier $(1 + E_n)^t$ has been called the reduction factor. Using this factor the expenditures made in different years can be reduced to the last year of the planning period. In the designated example it is essential not only to reduce the expenditures for the first 2 years under the first variation to the end of the second year but also for both variations to the end of the fifth year. In this instance all expenditures will be reduced to the same moment in time and, consequently, comparable from the viewpoint of the time of making the expenditures.

The Standard Procedure recommends that in the reduction formula use be made of a reduced effectiveness rate in comparison with the one employed in the formula for annual reduced expenditures. This is justified by the fact that not all the deferred amount of investments will be reput into economic circulation but only a certain portion corresponding to accumulation in national income. According to the 1977 Procedure for Determining the Economic Effectiveness of Using New Technology, the reduction rate for E has been set at 0.1.

Above we have examined the basic criteria for assessing the economic effectiveness of capital investments. Let us apply these to determining the economic effectiveness of BTS. A large technical system contains an enormous number of structural elements and possesses a complex hierarchical structure. Each element in the system is characterized by certain technical and economic parameters which, in turn, characterize its effectiveness. However the system's effectiveness is not the mere total of the effectivenesses of its elements. This stems from the important quality of large systems, that is, the presence of innate properties which do not derive from the known (observable) properties of the system's elements and the methods of connecting them. There is a new terminology and the essence of this characteristic property of systems was disclosed even by K. Marx who felt that "the force of attack by a cavalry squadron and the force of resistance of an infantry regiment are fundamentally different from the total of those forces of attack and resistance which the individual cavalry and infantrymen could develop.... Here it is a question not only of a rise by the cooperation of individual productive force but also one of creating a new productive force which in its very essence is a mass force."[2]

The parameters of the system elements and the system as a whole are interrelated and for this reason, along with the effectiveness criteria of the system's individual elements (the particular criteria) one must examine the criterion for the effectiveness of the system as a whole, or the general criterion.

Technical systems exist in time and space. The method of a system's existence in space can be described by models which are structural schemes of a system. In an analysis of large systems, ordinarily two types of models are used: canonical and hierarchical. A canonical model shows the functioning scheme of a system while a hierarchical model gives its structure and for this reason a system's canonical model can be used for establishing its general criterion while the hierarchical one is used for working out the hierarchy of particular criteria. Let us examine the questions of formulating the general criterion for estimating BTS economic effectiveness.

A system's canonical model reflects an aggregate of factors which characterize the process of its functioning through the external structure, the system's inputs and outputs. The inputs and outputs determine the relationships between the designated system and the environment. As is known, the inputs into a large system can be divided into: informational which define the system's operating program; resource

---

[2] K. Marx, "Capital," Vol 1, Politizdat, 1969, p 337.

which ensure the development, recovery and functioning of the system; the conditions and constraints imposed by interacting systems; the conditions and constraints imposed by national economic interests; the conditions and constraints imposed by nature. In an analogous manner the outputs out of the system are divided into informational which characterize the results of the system's informational activities; the consumption of resources as a result of the system's functioning; the conditions and constraints imposed by the results of the system's functioning on the interacting systems; the system's impact on the national economy; the system's impact on nature.

The listed inputs and outputs expressed by the corresponding parameters are an object of telic, operations and other types of research which in one form or another are generalized in the criteria. The individual inputs and outputs have a direct link with the nature of the criterion itself and with its structure. This applies first of all to the links which characterize the system's interaction with the national economy.

They mark first of all those constraints which are related to the influence of national economic interests on the system's functioning (for example, the constraints for the required resources and the functioning modes for which the systems are designed or can be designed) and the system's influence on national economic interests. Constraints can also be imposed on the system's functioning by nature (atmospheric conditions, terrain and so forth). In turn, the system itself influences the state of nature and the environment. Hence it can be concluded that the system estimate criterion should include a system of constraints from the national economy and the environment on the studied system. These constraints influence the system's technical and economic parameters.

The constraints imposed by the national economy on the resources required for the creation and functioning of the system or the constraints imposed by nature on the system's technical parameters (the noise level, the impurity content in exhaust gases and so forth) do not exhaust the influence of a system approach on shaping the BTS economic effectiveness criterion. No matter how complicated a technical system is, it, in turn, is an element of a higher level system.

The structure of these large systems is also hierarchical with a rather high level of hierarchy (system ranks). Previously, in Chapter 1, as an example we gave our nation's air transport system (a first-rank system) which consists of aircraft systems (second-rank systems) which in turn represent the air transport systems of the nation's economic regions. The second-rank aircraft systems can be divided into third- and fourth-rank systems and so forth.

Air transport, along with the other types of transport systems, comprises a higher rank system, the nation's unified transport system. Transport, along with industry, construction and agriculture, comprise the higher-rank system, the national economic system. Naturally in working out the effectiveness criterion for the aircraft system, it is essential to consider the influence of the given system's effectiveness on the functional effectiveness of the other systems which interact with this system on the scale of the national economic system. Such a requirement of systems analysis has been named the national economic approach to estimating the economic effectiveness of systems. This means that its adopted variation should be not only effective within that higher rank system in which the given system functions but also should contribute to increasing the effectiveness of the entire national economy.

The national economic approach to assessing the BTS is related primarily to its life cycle (NIR [scientific research work], OKR [prototype design work], series production and operation). A national economic systems approach to estimating the economic effectiveness of BTS is also based on considering the time and resource expenditures in the stages of the system's life cycle in the various spheres of material production: at the NII [scientific research institute], experimental enterprises, at serial-production plants and in operation.

In examining the stages of a system's life cycle, it must be stressed that operations are the determining stage as far as estimating a system's effectiveness. It is precisely this stage which discloses a system's specific consumer properties which determine its effect. Precisely at this stage it is essential to compare the effect from the system with the expenditures which include the expenditures on the stages of NIR, OKR and serial production. For confirming the given viewpoint, one can refer to the thesis of K. Marx on the economic limits of new technology under socialism as elaborated in "Capital." This thesis states: "If one views machines exclusively as a means for reducing the cost of a product, then the limit of their use is determined by the fact that the labor which their production costs should be less than the labor which is replaced by their use."[3] From this it can be concluded that the effect of the new technology is realized in the spheres where it is employed. The economic effectiveness of new technology must be calculated in these spheres.

Above we have examined the basic provisions in the theory of systems effectiveness and analysis and the theory of capital investment effectiveness which can be used in directly formulating the economic effectiveness criterion for the BTS. Proceeding from the adopted multicriterial approach to the estimate, we select minimum expenditures as the main criterion while the remaining criteria of the time and effect operate as the disciplining conditions (3.5).

The national economic approach makes it possible to determine the national economic effect of the BTS. This effect shows what benefit the national economy will obtain in using the new BTS. The national economic effect represents the total savings of expenditures for the entire aggregate of produced resources (labor, materials and fixed productive capital) as achieved by using the new system. Current expenditures are added to the one-shot capital ones by using the reduced expenditures formula. In comparing several variations of a system, the most effective one is chosen for the minimum reduced expenditures on operations:

$$C_r + E_n K_r = \min, \tag{3.16}$$

where C_r--annual current expenditures for operating the system;
K_r--capital investments for operating the system.

The expenditures on operating the system are dynamic. At the beginning of operations and then as the system is put into operation they decline. Moreover, the

---

[3] K. Marx and F. Engels, "Soch." [Works], Vol 23, p 404.

volume of the system's annual work also changes over time, and for this reason it is proposed that the average annual expenditures be considered for the entire operating period of the system $\overline{C}_r$. If we consider the average amount of expenditures for operations, then (3.16) assumes the form

$$\overline{C}_r + E_n K_r = \min. \tag{3.17}$$

Then it is essential to disclose the content of the concepts "average annual current expenditures for operations" and "capital investments into operations." Moreover, it is essential to bear in mind that according to the above-adopted national economic approach, in calculating BTS effectiveness one must not restrict oneself to current and capital expenditures just into operation. It is also essential to consider the current and capital expenditures into the NII, the prototype design organizations (OKO) and into experimental and serial-production plants.

The national economic expenditures for the system's entire life cycle includes the expenditures on scientific research Z_{nir}, design and prototype production work carried out at experimental enterprises Z_{okr}, expenditures on the serial manufacturing of the system Z_m and expenditures on operating the system Z_r. The dynamics of expenditures for the entire life cycle of a system can be shown schematically in the form of Fig. 3.2. The expenditures Z_{nir}, Z_{okr}, Z_m and Z_r include both the capital investments K_{nir}, K_{okr}, K_m and K_r as well as current expenditures C_{nir}, C_{okr}, C_m and C_r. All these expenditures are made at different times. Consequently, in order that the compared variations are compatible, all these expenditures must be reduced to the same moment in time. Within (3.17) let us try to consider the national economic expenditures and the time factor.

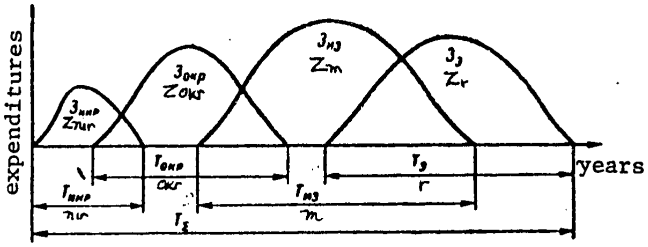


Fig. 3.2. Expenditure dynamics during life cycle of a system

As the reduction base for all expenditures, let us use the first year of operation. For the convenience of the calculations, we will assume that the end of the year is the point of reference. Thus, the reduction point will be the point $T_{sr}+1$, where T_{sr}--the point of starting operations in the interval T_Σ. Let us designate the reduction point by T_g. Let us introduce the other notations used below:

T_{snir}, T_{sokr}, T_{sm} and T_{sr}--the starting points for the stages of NIR, OKR, manufacturing and operating the system in the interval T_Σ (years); T_{scnir}, T_{scokr}, T_{scm}, T_{scr}--the starting points for carrying out capital investments for the corresponding stages; C_{enir}, T_{eokr}, T_{em}, T_{er}--correspondingly the ends of the same stages in the interval T_Σ (years); T_{ecnir}, T_{ecokr}, T_{ecm}, T_{ecr}--the ending points for carrying out capital investments for the corresponding stages; T_{nir}, T_{okr}, T_m, T_r--the duration of the stages for the NIR, OKR, manufacturing and operation of the system (years); C_{nirt}, C_{okrt}, C_{mt}, C_{rt}--the costs of NIR, OKR, manufacturing and operating the system in year t; K_{nirt}, K_{okrt}, K_{mt}, K_{rt}--capital investments in the corresponding spheres in year t.

The dynamics of the capital and current expenditures for the system's life cycle is shown in Fig. 3.3.

If one considers just the expenditures into operations, then considering the time factor, the comparative effectiveness indicator (3.17) assumes the form

$$Z = \frac{1}{T_r}\sum_{t=T_{sr}+1}^{T_{er}} C_{rt}(1+E)^{T_g-t} + E_n \sum_{t=T_{scr}+1}^{T_{ecr}} K_{rt}(1+E)^{T_g-t} = \min. \qquad (3.18)$$

In actuality (3.18) expresses the average annual reduced expenditures for operating the system. The variations are compared for the same amount of work and for the same operating period, and for this reason the total expenditures on operations reduced to the same moment of time, that is, to the start of operations, will be divided by the amount of the period T_r.

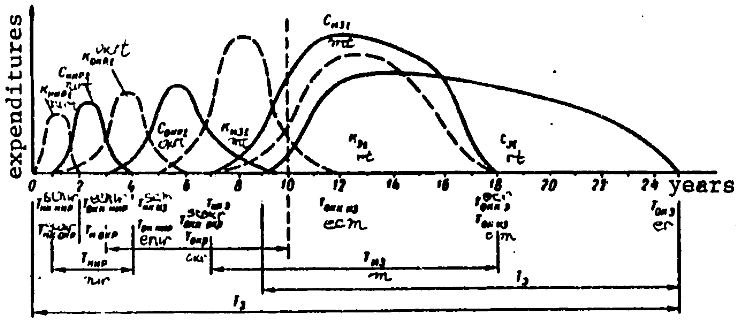


Fig. 3.3. Dynamics of capital and current expenditures over system life cycle

The capital investments into operation made during the various years using the reduction factor $(1+E)T_g-t$ also lead to the start of operations. Expression (3.18) also considers the expenditures on series production. The current operating expenditures C_{rt} consider the deductions for the renovation of the system while the amount of capital investments K_{rt} includes the cost of manufacturing the system. However, here no consideration is given to the current expenditures on NIR, OKR as well as the capital investments into NIR, OKR and serial production. The formula (3.18) also does not consider the capital investments into related production which must be carried out in line with the creation and operation of the system being evaluated K_{cp}. It can now be considered being generally recognized that all the listed expenditures must be considered in calculating the economic effectiveness of technical systems and they must be taken into account in the amount of capital investments.

However, before putting all the listed expenditures into (3.18), let us clarify this equation. Its first term $\frac{1}{T_r} \sum_{t=T_{sr}+1}^{T_{er}} C_{rt}(1+E)^{T_g-t}$ expresses the average annual outlays for operations during the operating period. These are average weighted expenditures for C_{ri} and these expenditures are weighted using the amount $(1+E)^{T_g-t}$--the reduction factor for expenditures made at different times. The expenditure reduction factor is also the weight which is considered in determining the average weighted expenditures. In determining the average weighted expenditure it is essential to multiply the elements of the aggregate by their weights and then divide this by the total of the weights. In this instance the total of the weights is $\sum_{t=T_{sr}+1}^{T_{er}} (1+E)^{T_g-t}$. Consequently, in (3.18) it is essential to divide the total of the products of the current expenditures and reduction factors by the amount $\sum_{t=T_{sr}+1}^{T_{er}} (1+E)^{T_g-t}$, and not by the operating period T_r. If this adjustment is incorporated in (3.18), then it will assume the form

$$Z = \frac{1}{\sum_{t=T_{sr}+1}^{T_{er}} (1+E)^{T_g-t}} \sum_{t=T_{sr}+1}^{T_{er}} C_r(1+E)^{T_g-t} + E_n \sum_{t=T_{scr}+1}^{T_{ecr}} K_r(1+E)^{T_g-t} = \min. \qquad (3.19)$$

After the adjustments made, let us draw attention to the second term of (3.19), to the capital investments made into the system's operations. Considering the national economic approach, it is essential to somewhat broaden the concept "capital investments into operation" and incorporate in it the cost of the NIR and OKR (NIOKR), the cost of the system and the cost of the capital investments into the production of the system and related production.

The cost of the system is divided into two components: the cost of the central elements in the system K_c and the cost of the other elements of the system K_v which ensure the functioning of the system's central elements. The central element has a determining impact on the parameters and functioning conditions of the other system elements, for example, the aircraft is the central element for an aircraft system.

Considering what has been examined above, it is possible to obtain the following formula for the reduced expenditures on operating the system:

$$Z = \frac{1}{\sum\limits_{t=T_{sr+1}}^{T_{er}} \alpha_t} \sum_{t=T_{sr+}}^{T_{er}} C_{rt}\alpha_t + E_n \left(\sum_{t=T_{scniokr+1}}^{T_{ecniokr}} K_{niokr} + \sum_{t=T_{niokr+1}}^{T_{eniokr}} C_{niokrt}\alpha_t + \right.$$

$$\left. + \sum_{t=T_{scr+1}}^{T_{ecr}} K_m\alpha_t + \sum_{t=T_{sccp+1}}^{T_{eccp}} K_{cpt}\alpha_t + \sum_{t=T_{scc+1}}^{T_{ecc}} K_{ct}\alpha_t + \sum_{t=T_{sco+1}}^{T_{eco}} K_{it}\alpha_t \right) = \text{min.} \tag{3.20}$$

where C_{rt}--annual current expenditures on operating the system;
α_t--reduction factor for expenditures at different times, $\alpha_t = (1+E)^{T_g-t}$;
K_{niokrt}--capital investments into NIOKR in year t;
C_{niokrt}--the costs of NIOKR in year t;
K_{mt}--capital investments into producing the system in year t;
K_{cpt}--capital investments into related production in year t;
T_{sccp}, T_{eccp}--beginning and end of making capital investments into related production;
K_{ct}--capital investments into operations needed for acquiring system central elements in year t;
T_{scc}, T_{ecc}--beginning and end of carrying out capital investments for acquiring system central elements;
K_{ot}--other capital investments into operation of system (excluding cost of system central elements) in year t;
T_{sco}, T_{eco}--beginning and end for carrying out other capital investments into operations.

Let us examine in more detail the content of the individual components of (3.20).

The annual current expenditures on operating the system C_{rt} consider all the expenditures for operating the system carried out in year t: expenditures on upkeep and maintenance, fuel and amortization deductions. Amortization deductions include deductions on major overhauls and deductions on the complete replacement of the system (renovation). As for the second term of (3.20), it includes the amount of the corresponding capital investments. In the early designing stages or even in the predesign stages, when the system's effectiveness is being calculated and when the system's effectiveness is being forecast, the question arises of the methods to determine this amount. Many works determine capital investments from a scheme of reduced expenditures. Let us given certain examples.

The 1964 "Basic Methodological Provisions on Determining the Economic Effectiveness of Scientific Research" recommend that economic potential be determined by the formula

$$Y_g = (C_1-C_2) + E_n(K_1-K_2) + T_n'(C_1'-C_2') + (K_1'-K_2'),$$

where C_1 and C_2--the production costs of the annual product volume of the sector producing the means of labor, respectively, for the new and base variation;

K_1 and K_2--capital expenditures needed for the above-indicated product output in the sector producing the means of labor, respectively, for the new and base variations;
C_1' and C_2'--the cost of the annual volume of the product produced with the designated means of labor in the sector using them as fixed capital;
K_1' and K_2'--additional capital expenditures (in addition to the expenditures on the designated means of labor) in the sector using these means of labor as fixed capital;
E_n--the normed effectiveness coefficient in the sector producing the means of labor;
T_n'--the normed repayment time in the sector using the designated means of labor as fixed capital.

In this expression the reduced expenditures are used for determining the value of the means of labor.

The 1972 "Instructions on Determining the Economic Effectiveness of Capital Investments in Construction" employ an indicator for the full reduced expenditures and these consider the expenditures in creating the project and the expenditures for operating it. The full reduced expenditures are calculated by the formula:

$$P_i = C_i + E_n K_i + E_n K_i' + M_i T,$$

where C_i--the cost of construction-installation work in building a project using variation i;
K_i--capital investments into the fixed productive capital and investments into working capital in the construction sphere;
K_i'--related capital investments into producing building materials and structural elements;
M_i--operating expenditures (average annual);
T--the calculation period during which the operating expenditures are considered (this can be set equal to the normed capital investment repayment time).

In this formula as well the expression $C_i + E_n K_i$ is used to determine the cost of the project. It would be possible to give a whole series of other works in which reduced expenditures are interpreted as the price or value. This position has been expressed most clearly in the work [34, p 53]: "Economic effectiveness in the national economy is ultimately measured by the growth of social labor productivity which is determined either by the quantity of manufactured product per unit of time or by the expenditures of live and embodied labor per article or per unit of work (the cost of the article or a unit of work). The lower the cost of the article (unit of work) the higher the labor productivity.

"In the sphere of aircraft system operations a standard operation can be accepted as a unit of work and, consequently, the cost of the operation is an indicator of social labor productivity. The reduced expenditures also form the criterion for the operations' cost."

According to [34] in a general form the criterion for the cost of an operation can be expressed in the following manner:

$$\tilde{C}_{op} = \int_{\tau_0}^{\tau_K} \sum_{i=1}^{n} \sum_{j=1}^{m} \tilde{C}_{1ij}(\tau) n_{ij}(\tau) d\tau, \tag{3.21}$$

where $\tilde{C}_{1ij}(\tau)$--expenditures per unit of means per unit of time τ;
$n_{ij}(\tau)$--the quantity of means required per unit of time τ;
i--the index of the type of aircraft participating in the operation, $i \in [1; n]$;
j--the index of local operations from which the global operation is formed, $j \in [1, m]$;
τ_0 and τ_K--the time for the beginning and end of the operation.

In the simplest case the formula (3.21) assumes the form

$$\tilde{C}_{op} = \tilde{C}_1 n.$$

Expenditures per unit of means $\tilde{C}_1$ are determined from the formula

$$\tilde{C}_1 = C_1 + E_n K_{ydr}, \tag{3.22}$$

where C_1--the cost of a unit of means;
K_{ydr}--the proportional, that is, per unit of means, capital expenditures.

Let us adopt the reduced expenditures for determining the amount of capital investments in (3.20). The capital investments into prototype or serial production can be represented as the total of the annual reduced expenditures on building one or another project. Analogously it is possible to represent capital investments into related production and operation (minus the cost of the system central elements). The cost of construction can be expressed by the corresponding reduced expenditures of the construction organization. However in system feasibility studies, as a rule, there is no need to determine the cost by such a differentiated manner, as there are determined proportional capital investments which are used in the calculations. The organizations which work out the norms for proportional capital investments in determining capital investment costs use the reduced expenditure scheme. Let us examine in more detail the procedure for determining the cost of the central elements in a system. This can be represented as the reduced expenditures in year t

$$Z_{mt} = C_{mt} + E_n K_{mt}. \tag{3.23}$$

The operating capital investments needed for acquiring the system central elements K_{ct} can be represented in the following form:

$$\sum_{t=T_{scc}+1}^{T_{ecc}} K_{ct} = \sum_{t=T_{sm}+1}^{T_{em}} Z_{mt}.$$

If the cost of the system central elements is to be determined using the reduced expenditures, then the amount of corresponding capital investments can be expressed by the formula

$$\sum_{t=T_{sm}+1}^{T_{em}} Z_{mt} = \sum_{t=T_{sm}+1}^{T_{em}} (C_{mt} + E_n K_{mt}).$$

Formula (3.23) needs a certain adjustment. First of all this applies to the concept of K_{mt}. It designates the new production capital which embodies the necessary capital investments. However, in determining the value of the system's elements, one must not restrict oneself to just the new capital and existing capital must be considered as well. For confirmation of this let us examine the following example.

Let us assume that according to one variation of a system capital investments are required into serial production, while under different variations they are not needed, that is, the elements of the system are manufactured at an operating enterprise without any reconstruction. If in determining the cost of the system we consider only the capital investments and do not consider the value of the existing capital (let us remember that here we do not have in mind the amortization deductions), then according to the second variation the system's cost (if it is figured from the above-given scheme of reduced expenditures) will equal the system's cost as capital investments equal zero. However from the operational viewpoint the capital investments for both variations should be considered not in terms of production costs but rather in terms of value. For this it is obviously essential to consider not only the capital investments but also the existing capital. If the amount of the existing and new capital is considered, then the cost of the system's elements is determined by the formula:

$$\sum_{t=T_{sm}+1}^{T_{em}} C_{mt} + E_n \sum_{t=T_{sm}+1}^{T_{em}} (F_{mt} + K^n_{mt}),$$

where F_{mt}--the amount of existing capital used for producing the system's elements in in year t;

K^n_{mt}--the amount of new capital in year t (capital investments into production in a running total up to year t inclusively which are used in year t in manufacturing the system's elements).

Let us make one other clarification. The operating capital investments needed for acquiring the system central elements do not equal the cost of manufacturing these elements over the entire period of their production. In the amount of the capital investments K_c it is essential to consider only that value of the system elements which ensures expanded reproduction in operation. The capital investments going to replace the wear on the system's elements, that is, used in simple reproduction, must not be considered in the amount K_c. This means that expenditures in production must be totaled for the production period of the system's elements going to expand the capital in operation from the start of manufacturing $t = T_{sm}+1$ to $t = T_{em}$, that is, to the moment when the elements are used to replace the worn out.

Hence the amount K_c is determined from the following formula:

$$K_c = \sum_{j=1}^{m} \sum_{t=T_{jsm}+1}^{T^{T}_{jem}} Z_{jmt},$$

where j--system element;
m--number of system elements.

If the remaining expenditures for the system are represented as the total expenditures for its elements, then formula (3.20) assumes the form:

$$Z = \frac{1}{\sum\limits_{t=T_{sr}+1}^{T_{er}} t} \sum_{t=T_{sr}+1}^{T_{er}} C_{rt}\alpha_t + E_n \sum_{j=1}^{m} \left(\sum_{t=T_{jscniokr}+1}^{T_{jecniokr}} K_{jniokrt}\alpha_t + \sum_{t=T_{jniokr}+1}^{T_{jeniokr}} C_{jniokrt}\alpha_t + \right.$$

$$\left. + \sum_{t=T_{jscm}+1}^{T_{jecm}} K_{jmt}\alpha_t + \sum_{t=T_{jsccp}+1}^{T_{jeccp}} K_{jcpt}\alpha_t + \sum_{t=T_{jsm}+1}^{T^{T}_{jem}} Z_{jmt}\alpha_t + \sum_{t=T_{jsco}+1}^{T_{jeco}} K_{jot}\alpha_t \right) = \text{min.} \qquad (3.24)$$

The designated criterion for the national economic effectiveness of the BTS (3.24) makes it possible to select the most effective variation of the system. Its economic essence is that it expresses the value of the result from the system's functioning (the value of its work or the value of its carrying out of the set operation).

The national economic approach to determining BTS effectiveness does not exclude the elucidation of cost accounting results or a cost accounting effect but, on the contrary, in taking a decision on the development or introduction of a new BTS, one must not restrict oneself to disclosing the cost accounting effect and it is essential to also consider the national economic effect. Let us illustrate this idea by the following hypothetical example. Let us assume that an aircraft designer has taken a decision to increase its service life by (hours). The expenditures of the experimental enterprise to increase the aircraft's service life are 1 million rubles. These expenditures are financed from the budget and are not incorporated in the aircraft price. As a result of increaseing the aircraft service life, civil aviation gains a savings totaling 50,000 rubles a year in amortizing the aircraft. This comprises the annual cost accounting effect of civil aviation. From the viewpoint of the civil aviation's cost accounting interests, an increase in aircraft operating life is advisable. However, from the viewpoint of the national economy where it is also essential to consider the expenditures on phototype design work to increase the aircraft service life, this measure is not advantageous. The additional 1 million rubles of expenditures on the OKR will be paid back in 20 years in civil aviation (1 million rubles : 50 rubles). Consequently, proceeding from national economic interests, the OKO must seek out more economic means for increasing aircraft life.

The essence and relationship of the national economic and cost accounting effects of new technology has been disclosed most completely in the work of L. M. Gatovskiy [9]. The national economic effect makes it possible to answer the question of the economic advisability of new technology for the national economy as a whole. The category of the cost accounting effect provides an answer to the question of the cost accounting advisability for the manufacturers to produce and for the consumers to employ a given new technology which is effective for the national economy. Thus, the national economic effect is primary and the cost accounting effect is secondary. The task is not to set one effect in opposition to the other but rather achieve their unity with the determining role for the national economic effect in relation to the cost accounting one. For this the socialist state uses the economic mechanisms and levers achieving an optimum reconciliation of the interests of society and the individual enterprises.

The interaction between the national economic and cost accounting effects occurs in the following sequence: the national economic effect--selection of new technology--price of new technology--cost accounting effect.

A clear differentiation of the two effects--national economic and cost accounting--has a positive impact on the further development of the theory of new technology effectiveness. If this idea becomes generally accepted, then the supporters of either types of effect or either type of accounting will debate not the replacing of one accounting by another or one effect by another but will jointly seek out ways for the reconciling and interaction of these two types of effect. As is known, for a long time there have been disputes between the supporters of a uniform capital investment effectiveness norm for the national economy as a whole and the supporters of differentiated rates for the sectors. We feel that it would be possible to provide the following reconciliation of these norms: in calculating the national economic effect to employ a unified national economic norm for capital investment effectiveness and in calculating the cost accounting effects the sectorial norms for capital investment effectiveness. If a uniform normed effectiveness coefficient would be approved in setting the sectorial procedural instructions on determining economic effectiveness for the use of new technology for all the sectors then the national economic and cost accounting effectivenesses would be figured using a uniform norm.

The cost accounting effectiveness is based upon the mechanisms of planning, expenditure accounting, price formation and encouraging new technology. With the existing planning and economic incentive system, as is known, profit and profitability are the important cost accounting indicators for the effectiveness of the sectors and enterprises. For this reason, for the large technical systems in national economic use and which are financed from centralized capital investments, the criterion of cost accounting effectiveness is formed using the profit and profitability indicators.

According to the 1977 Procedure for Determining Economic Effectiveness of New Technology, the cost accounting effect from the production and use of new technology is determined by the formula:

$$Y_x = \sum \Delta P - E_n \sum \Delta K, \qquad (3.25)$$

where Y_x--the aggregate cost accounting effect for production from the output and use of new technology in the year being planned t, rubles;
$\sum \Delta P$--increase in profit (decline in cost) from all measures under the new technology plan in the year being planned t, rubles;
$\sum \Delta K$--capital investments for all measures under the new technology plan of year t, rubles.

This indicator is used in determining the cost accounting effect of the new BTS, and for this it is essential to compare the increased profit in the set period ΔP with the normed amount of effectiveness from the additional capital investments $E_n \Delta K$ using the formula

$$Y_x = \Delta P - E_n \Delta K. \tag{3.26}$$

The indicator of (3.26) makes it possible to determine the overall cost accounting effect of the new system. The new system's effectiveness level is determined by the effectiveness indicator for additional capital investments:

$$E_p = \Delta P / \Delta K.$$

Along with the comparative effectiveness indicator it is also recommended that the general effectiveness indicator be used, that is, the indicator for the effectiveness of full capital investments

$$Y_{cc} = P/K,$$

where P--annual profit.

The economic effectiveness criteria for the BTS make it possible to determine the economic effect from its creation and introduction. First of all it is essential to distinguish three types of economic effect: preliminary, expected or planned and actual.

In the stage of carrying out scientific research for establishing the necessity of the OKR, that is, before taking the decision to carry out the OKR to develop a new BTS, the preliminary economic effect should be determined.

After completing the development of the new system, the expected economic effect is determined considering the development results and the assumed areas and conditions for its introduction.

At the stage of drawing up the new BTS introduction plan, the expected economic effect is the planned effect.

After the new system has been put into operation the actual economic effect is determined.

Each of the above-designated types of effect (national economic and cost accounting) can be determined as preliminary, expected or planned and actual. If one generalizes the characteristics of the national economic and cost accounting effects, it is possible to emphasize that both types of effect are figured proceeding from common methodological principles. The differences are in the nature of the initial data, methodological principles. The differences are in the nature of the initial data, the normative materials and the comparison base for elucidating the effect.

The nature of the initial data, their completeness and degree of certainty depend, in the first place, upon the life cycle stage of the BTS for which its economic effectiveness was figured and, secondly, upon the level on which the problem is solved (the national economy or sector). In determining the annual economic effect of a BTS, the following are used as the base of comparison:

1) During the stage of formulating the NIOKR plans (in the process of selecting the variation for creating a new BTS), and in taking a decision to put a new BTS into production--the indicators of the best BTS designed in the USSR (or a foreign BTS which can be purchased or developed in the USSR on the basis of licensing); in the instance of the absence of design studies in the USSR and the impossibility of using foreign experience as a basis of comparison, the indicators of the best BTS in the USSR are used;

2) In the stage of formulating the plans for the development of the new BTS as well as in the stage of their introduction and operation--the indicators of the BTS to be replaced.

In addition to the above-mentioned types of effect from the BTS, depending upon the adopted calculation period, a distinction must be made between the annual economic effect of a new BTS Y_g and the complete economic effect of a new BTS over the period of its operation Y . The annual national economic effect from operating a BTS is calculated on the basis of the criterion of (3.24) using the formula

$$Y_g = Z_b - Z_n, \tag{3.27}$$

where Z_b and Z_n--the average annual reduced expenditures for the development and operation of the new and base systems.

The full economic effect over the operating period is calculated by the formula

$$Y_\Sigma = (Z_b - Z_n) \sum_{t=T_{sr}+1}^{T_{er}} \alpha_t. \tag{3.28}$$

Along with the absolute amount of the new effect for a new technical system (the annual or full effect), the level of its effectiveness is also determined

$$E_p = \frac{\bar{C}_{rb} - \bar{C}_{rn}}{K_{\Sigma n} - K_{\Sigma b}}, \tag{3.29}$$

where $\bar{C}_{rb}$, $\bar{C}_{rn}$--the average annual cost of operating the base and new variations of the BTS;

$K_{\Sigma b}$, $K_{\Sigma n}$--the total capital investments for the development, production and operation of the system reduced to the start of the system's operation.

CHAPTER 4: FORECASTING THE COST ESTIMATES OF LARGE TECHNICAL SYSTEMS

Economic forecasting of BTS development assumes the choice of the most effective areas of scientific and technical development using economic effectiveness criteria. Such a choice becomes objective only under the condition that the possible development alternatives are coordinated with the estimates for resources needed to implement these alternatives.

The resources which ensure the scientific and technical development of the BTS are diverse, but, as a rule, can be put in four basic groups: labor, material, energy and information. The direct totaling of the needs for each of the resource varieties included in these groups is excluded as the resources are incommensurable in terms of their natural and physical form. This circumstance forces one to resort to cost estimates of resources which to a significant degree reflect the expenditures of social labor included in all resources without exception, independently of their nature and physical embodiment. In addition to ensuring the internal comparability of the resources, the cost estimates are a means for comeasuring the amount of resources with the end effect (the result) or BTS development to the degree that the effect of any development can be measured by cost estimates.

The cost estimates for resources which ensure the development of the BTS are concretely expressed in the value of the scientific research and prototype design work, serial production and operation. In the book these are linked together by the common concept of the BTS cost estimate.

4.1. Forming the BTS Cost Estimates

The complexity and great scale of the BTS lead to the spreading out of the process of their creation (NIOKR) and serial production in many industrial sectors. Each sector, and within the sector the individual NII, OKB and serial-production plants are specialized in the creation and industrial output of certain groups of the BTS functional elements. These elements are united by the common purpose and character of the physicochemical functioning processes, by designing and production processes. The particular features of the individual functional elements of the BTS are reflected in the organization and content of the processes of their creation and serial production, in the composition of the consumed resources and in the structure of the cost estimates. All of this ultimately leaves an impression on the formative patterns in the cost of creating and serially producing related groups of BTS functional elements. An analogous situation is also observed in operations.

Thus, for each group of related elements there is a corresponding special inner logic and structure for the processes of NIOKR, series production and operation. Because of this the entire process of forming the BTS cost estimates can be broken down into a multiplicity of particular processes which to a significant degree occur independently.

In taking into account what has been said, the model of the BTS cost estimates which most reflects the process of their formation can be represented by the expression

$$C_{\Sigma} = \sum_{i}^{n} \sum_{j}^{m} C_{ij} n_j + \sum_{i} C_{s_{oi}}, \qquad (4.1)$$

where C_{ij}--production costs of system's element j at stage i of the life cycle;
n_j--number of j elements in system;
C_{soi}--the cost of the work related to organizing the system at stage i of the life cycle.

Such an interpretation of the process of forming the cost estimates of the BTS possesses two basic merits. In the first place, it considers the particular features of the functional elements and the conditions of their creation, serial production and operation. This provides a high sensitivity of the BTS cost estimates to a change in the characteristics of the individual elements and the stages of their life cycle. Secondly, the model (4.1) possesses great flexibility as an economic forecasting tool as it meets certain specific demands imposed by the logic of the BTS development process.

In actuality, in accord with the BTS development goals, the properties of the individual functional elements do not change uniformly. The increase in the system's functional possibilities in the individual stages is brought about by the development of those elements which at the given moment most impede overall development and thereby are the main obstacle on the path to achieving the general goal. Here the contribution of a specific element to overall development can vary in rather significant limits. But the demand for resources ensuring the development of the given element does not always correspond to this contribution.

Thus, if the element's functional parameters are close to their theoretical limits, an insignificant increase in them, nevertheless, can require significant expenditures. Consequently, even a small increase in an individual element's functional parameters can cause an increase in expenditures which is very tangible on a scale of the entire system. On the other hand, there are contradictory links between the parameters of the system's individual elements and for this reason in the process of developing the BTS, as a rule, compromises are made with a certain parameter being improved at the expense of a certain deterioration in the other parameters. Here as a whole the system's effectiveness rises. Thus, a change in the BTS cost estimates will depend upon the particular features of the development process of the individual elements which are characteristic for the given period of the system's development.

From what has been said it follows that a study of the formative process of the BTS cost estimates should start by an analysis of the system's internal structure and the isolating of individual elements or kindred groups into an independent research object.

The cost estimates for the functional elements are by nature dynamic.

In the process of developing the BTS functional elements, the expenditures at each stage of their life cycle are influenced by a complicated combination of factors of the most diverse nature. The influence of some is limited to an individual stage of the life cycle while others encompass all stages; the action of one is discrete while that of others is continuous. A majority of factors is in a complex interaction.

Fig. 4.1 gives a classification based upon the principle of dividing the total number of factors into related groups in accord with the place, time and nature of their influence on the cost estimates. As is seen from the diagram, all factors are divided into three large groups: factors characterizing the conditions of the creation, production and operation of the elements or the organizational-economic ones; the factors characterizing the functional element as a means of attaining certain goals or the operational effectiveness factors; the factors characterizing the BTS element as an object of creation, production and operation, or the design-production ones. Such a classification makes it possible:

1) To unify in each group the similar factors reflecting the influence of definite processes related to the development of the BTS on the cost estimate;

2) To utilize the link between the factors existing in each group and between the groups for the purpose of choosing the indicators which accumulate the influence of the factors related by cause-and-effect ties;

3) To disclose the mechanism of influence of the factors in each group on the cost estimates.

The operational effectiveness factors hold a central place in research on the formative patterns of the cost estimates, since precisely this group of factors accumulates the influence from the properties and functional capabilities of the BTS elements on the expenditures. In this group of factors it is possible to isolate two independent subgroups of which the first describes the element's ability to perform the set functions, or the functional subgroup, and the second is the maintaining of this capacity in the process of operation, or the technical and operational subgroup.

In the first subgroup are such factors as the element's purpose, the number of performed functions, the number of functional cycles within one operation, the maximum possible values of the functional parameters, the range of their change and so forth. The maintaining of the capacity to perform the set functions is characterized by reliability, service life and the repeated use of the functional element.

The influence of operational effectiveness factors on the cost estimates is manifested through the design and production features of the elements, including:

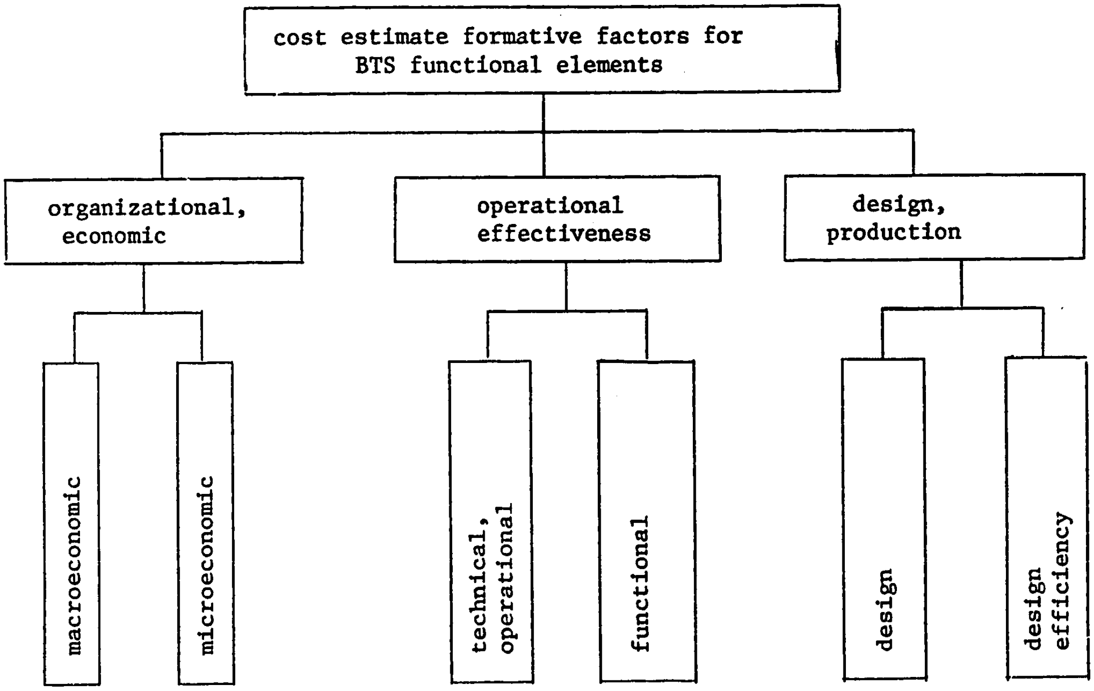


Fig. 4.1. Classification of formative factors for the cost estimates of BTS functional elements

weight, geometric shape, layout, the total number of design elements and their density in a unit of volume, the mechanical, physical and chemical properties of the employed materials; the manufacturing precision for the design elements and their connections; the roughness of the worked surfaces and so forth. The more complex the design of the BTS functional element and the higher the properties of the employed materials and the demands made on the quality of the production processes the more expensive it costs to create, produce and operate the element.

However, the mechanism of the influence of the operational effectiveness factors on the cost estimates is much more complicated as in each specific instance there is a multiplicity of possible combinations for the forms and methods of materializing the functional and technical-operational characteristics.

At the same time, the amount of the cost estimates for the elements, with other conditions being equal, can change depending upon the overall ratio of the number of original, standardized and unitized parts used in its design. A high degree of standardization and unitization, in deepening specialization and creating favorable conditions for the organizing of mass production for the individual design elements, leads to a marked reduction in the cost estimates.

The cost estimates of a new functional element will differ from the cost estimates of its predecessors the less the higher the degree of its design succession, that is, the overall correlation between the borrowed and newly developed design elements. The cost estimates will be influenced by the methods of assembling the designs into individual units, panels and assemblies, their interchangeability and the possibility of disassembly and assembly without disrupting the overall arrangement of the elements in the system. In other words, the cost estimates will depend upon the design efficiency of the functional element. Considering what has been said, it is necessary to isolate an independent group of factors, the design-technological ones, with the splitting of these into two subgroups: design and design efficiency.

The degree to which the designated factors influence the cost estimates of the BTS function elements depends largely upon the characteristics of the state of the processes occurring in the individual stages of the element's life cycle. These characteristics in the adopted classification are part of the microeconomic factors, that is, the factors operating on the enterprise and organization level. These include the organizational-technical level, the scale and degree of development of the processes involved in the creation, production and operation of the systems. These factors will be examined in detail in the subsequent sections of the chapter as for each stage in the life cycle their influence has its specific features.

Let us stop in somewhat greater detail on the influence of the organizational-technical level of creating and producing the systems on their cost estimates to the degree that this is related to the overall process of scientific and technical development of the BTS. The organizational-technical level of creating and producing the elements of the systems is characterized by the following basic indicators: the level and progressiveness of the production-technical facilities; the technical- and energy-to-labor ratio; the level of automating the mechanizing the labor processes; the progressiveness of the employed production processes, the level of technological outfitting; the system of preparing, servicing, controlling and planning production; the forms of the organization of labor and production; the level of concentration, specialization and cooperation [subcontracting].

The process of improving the technology and the technological and organizational forms for creating the producing the systems is an objective expression of scientific and technical progress. In essence, this process, in being a response to the development of the systems, makes possible the materialization of the new functional properties of their elements. For this reason the processes involved in the development of the technical and organizational-economic systems occur in parallel. Here there can also be certain temporary breaks in the pace of this development. The cost estimates simultaneously accumulate the influence of these two trends and respond to changes in the ratio of the development rates. There is no doubt that this important aspect in the process of forming the cost estimates should be taken into account in working out the cost forecasts.

Nevertheless, the present methods of forecasting the cost estimates proceed from the assumption that the one-time set development rates of the technical systems and the means ensuring this development remain constant. This is tantamount to the assumption that the technical and organizational-economic systems are in a state of relative equilibrium. Let us endeavor to analyze this question and for this we will turn back to the already-discussed particular features of BTS development.

The development of structural elements in systems occurs unevenly. One part of the elements undergoes substantial changes, a second changes insignificantly or not at all while the characteristics of a third can even deteriorate somewhat. For this reason at each moment of time the means and methods for ensuring the new properties of elements will comprise a certain portion in the aggregate of all the means and methods employed on a system-wide scale. It is also essential to consider that a portion of the parts and assemblies of the functional elements, in undergoing change, has been borrowed from previous designs.

Thus, at each stage in the development of a system a certain set of tested means and methods will be employed to ensure the functional possibilities of the system. And these means and methods will be constantly improved due to the continuous nature of the development processes and the action of the law of increased social labor productivity. The latter is an objective prerequisite for reducing the cost estimates. Consequently, the more intense growth of the cost estimates caused by certain breaks in the development pace of the individual BTS elements and the organizational-economic systems will to a significant degree be compensated for by a drop in the cost estimates of the individual BTS elements as well as the borrowed design elements.

The formation of the cost estimates is also influenced by factors operating on the national economic and sectorial levels, or macroeconomic factors. These factors comprise a group of conditions inherent to a certain economic development stage on the national scale. This includes the level of social labor productivity, the adopted forms of wages and incentives for the participants of material production, the system of relationships between the individual enterprises and the state budget, price policy and so forth.

Social labor productivity is the most important factor determining the overall level of the cost estimates.

To the degree that the rise in social labor productivity is expressed in a savings of both live and embodied labor, it works to reduce product costs. The continuous rise of social labor productivity and the decline in product costs are an objective reality determined by scientific-technical progress, by the rise in personnel skills and by the improvement in technology and the forms of industrial production.

The effect from the increase in social labor productivity is apparent in all stages of the BTS life cycle through the expenditures of past labor consumed in the process of creating, producing and operating the systems and embodied in the raw products, materials, fuel and means of production. Social labor productivity has a particular effect on the cost estimates of functional elements the designs of which employ new still not sufficiently developed materials, for example, titanium alloys (Fig. 4.2) or filamentary crystals (Fig. 4.3).

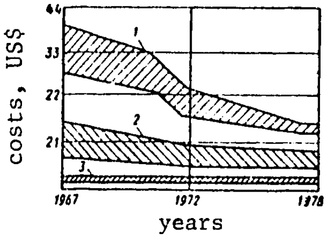


Fig. 4.2. Forecast of comparative cost of titanium sheet 1, titanium new plates, strips and rod 2, steel and aluminum 3

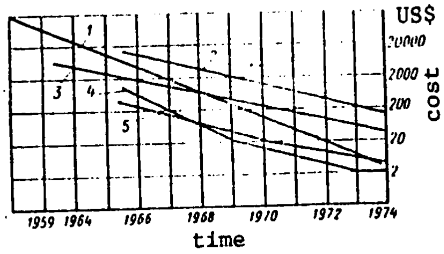


Fig.4.3. Change in the cost of certain fibrous crystal materials in the process of developing their production: 1--Filamentary sapphire crystals; 2--Fibrous silicon carbide; 3--Filamentary boron crystals; 4--Graphite fiber; 5--Filamentary silicon carbide fibers

At the same time, in analyzing, for example, the drop in the cost of materials made from titanium over time, it is not hard to see that as the new materials are developed and they become ordinary ones, this channel for the influence of social labor productivity on the cost estimates weakens. Thus, the intensive drop in the cost of titanium observable in the period from 1967 through 1972 was gradually replaced by a gentler dynamic so that the rate of decline in outlays on titanium materials became comparable with the trend for the change in the cost of steel and aluminum.

The macro- and microeconomic processes are processes with both a direct link and feedback. Thus, the satisfying of the new demands made on the organizational and technical level of the BTS production sectors determines the areas for the scientific and technical development of the organizational-economic systems and this ultimately leads to macroeconomic shifts and, in particular, to a rise in social labor productivity. This is tantamount to the saving of social labor. The latter, in turn, determines the amount of resources which are to be allocated to satisfy the social needs in the BTS and are to be channeled into their scientific and technical development. The interaction of the different-natured factors and their relation to the cost estimates are shown schematically in Fig. 4.4.

The dynamics of the factors in each group forms its own individual time trend for the BTS cost estimates in the process of their technical and scientific development while the entire aggregate of factors forms the resulting trend. Fig. 4.5 hypothetically shows the formation of the resulting trend of the cost estimates for the BTS elements against the background of the growth of their functional parameters. Each of the presented curves depicts an individual trend of the cost estimates formed by the dynamics of an individual group of factors when the remainder are conditionally assumed to be constant. The straight line parallel to the time axis depicts the hypothetical trend of the cost estimates J_c for the case where the factors involved in forming the cost estimates remain constant.[1]

As is seen from Fig. 4.5, the curve for the resulting trend changes direction: the decline observable at the start of the system's development is replaced by a rise in the cost estimates as the intensive development phase approaches. This occurs because in the first stages of the system's development the factors relating to the starting up of the processes of prototype and serial production of the functional elements are very strong and for this reason the prototypes for the functional

---

[1] Fig. 4.5 considers the influence of factors which have a steady time trend. However the cost estimates are also influenced by factors the changes in which can be step-like, cyclical, pulse or simply of a random nature. Thus, the real trends in the cost estimates are much more complex.

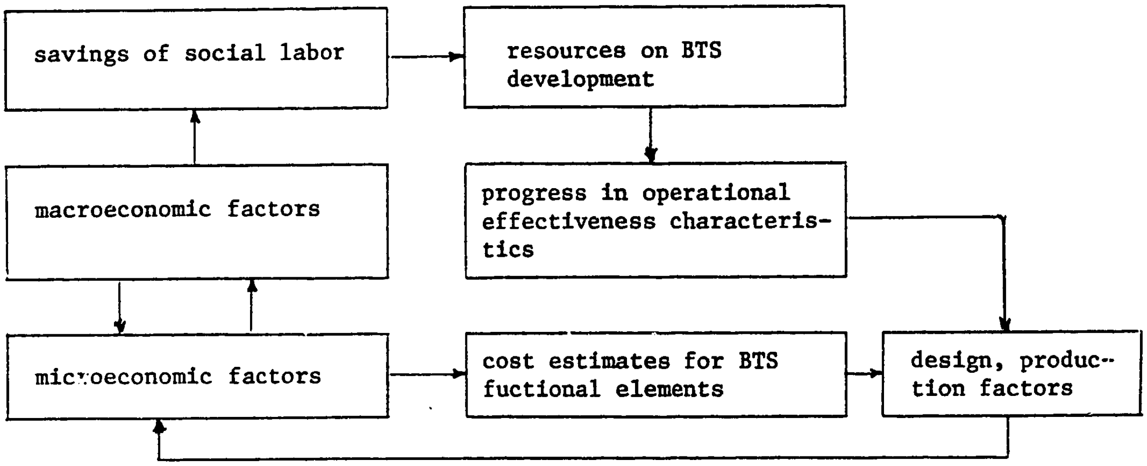


Fig. 4.4. Schematic diagram for interaction of factors which determine the cost estimates for the BTS functional elements

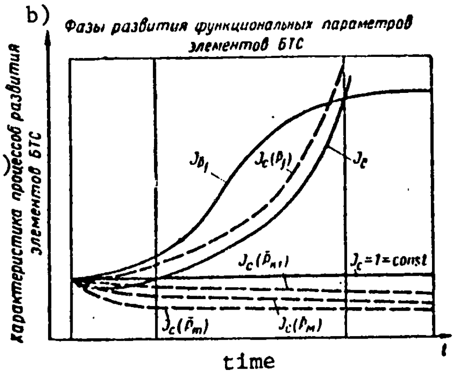


Fig. 4.5. Hypothetical scheme for forming the resulting trend $J_{\bar{c}}$ for the cost estimates of the BTS elements:

$J_c(\bar{P}_f)$, $J_c(\overline{P_{KT}})$, $J_c(\bar{P}_m)$, $J_c(\bar{P}_M)$--indexes of cost estimates determined, respectively, by dynamics of functional parameters, by the characteristics of the design efficiency, by micro- and macroeconomic factors; JP_f--index of functional parameters; a--characteristics of development processes for BTS elements; b--development phases of functional parameters for BTS elements

elements of the new class of systems, with other conditions being equal, have a higher production cost due to the newness of the designs and the lack of sufficiently developed production methods and practical skills for manufacturing them. Subsequently the influence of the designated factors weakens and the growth factors of the functional parameters and the steady trends in the macro- and microeconomic factors and design efficiency performance become the determining ones.

Thus, it can be stated that an improvement in the operational effectiveness characteristics determines the growth of the cost estimates while the progress in organizational-economic factors and design efficiency performance helps to reduce the cost estimates, however the resulting trend is a rising one.

A review of the general patterns in the formation of BTS cost estimates makes it possible to represent a model of the cost estimates in the following manner:

$$C = \sum_{i=1}^{n} \sum_{j=1}^{m} n_j C_{ij} (\bar{P}_f, \bar{P}_M, \bar{P}_m, \bar{P}_{KT}, t) + \sum C_{S0i}, \tag{4.2}$$

where $\bar{P}_f$--the vector for the functional and technical-operational parameters of the element;
$\bar{P}_M$--vector of macroeconomic variables;
$\bar{P}_m$--vector of microeconomic variables;
$\bar{P}_{KT}$--vector for design efficiency performance;
t--time elapsed since start of system's development.

From all that has been said it follows that in the most general instance a cost forecast should be based on scientific-technical, socioeconomic and organizational-economic forecasts. Considering that none of these forecasts can be compiled in isolation from the others, as the effective factors are closely interlinked, the cost forecast cannot be absolute. In other words it is possible to draw up a cost forecast which considers changes in the operational effectiveness parameters only for a completely definite combination of vectors which accumulate the influence of the other factors, since each new combination of factors will determine its own amount of the cost estimate.

Thus the cost estimate represents information of the following sort: event A can occur with a certain probability, if event B will occur under the condition that the occurrence of event B is accompanied by the simultaneous appearance of events C, D, E and so forth. Since the number of combinations for the simultaneous satisfaction of a multiplicity of conditions is incalculable, it is not hard to imagine that very high degree of ambiguity which is encountered in forecasting value estimates even for a short period of time. This ambiguity can be reduced only by methods based upon a knowledge of the patterns in the formation of the cost estimates during the developmental prehistory of the cost estimate object. In turn, the establishing of firm patterns becomes possible with complete systematized information on the processes of the object's scientific and technical development and actual resource consumption related to this development. However, the economic forecasting tasks arise in the most different stages of BTS development, and for this reason in forecasting one must frequently be content with information acquired up to the time of the rise of the problem. The volume and nature of the initial information also determine ultimately the existing differentiation of forecasting methods for the cost estimates.

4.2. Basic Principles and Methods for Forecasting Cost Estimates

4.2.1. The Choice of a Forecasting Method for Cost Estimates

The methods of forecasting cost estimates from the viewpoint of the principles of forecast information generation can be divided into two large groups. In the first group are the methods which are based upon the use of expert judgments for forecating purposes, that is, the expert estimate method, while the second group includes forecasting methods where the basic source of forecast information is a model for the formation of cost estimates or interpolation and extrapolation methods (Fig. 4.6). It must be pointed out that the involvement of experts in the compiling of cost forecasts is very useful when the second group of methods is to be used.

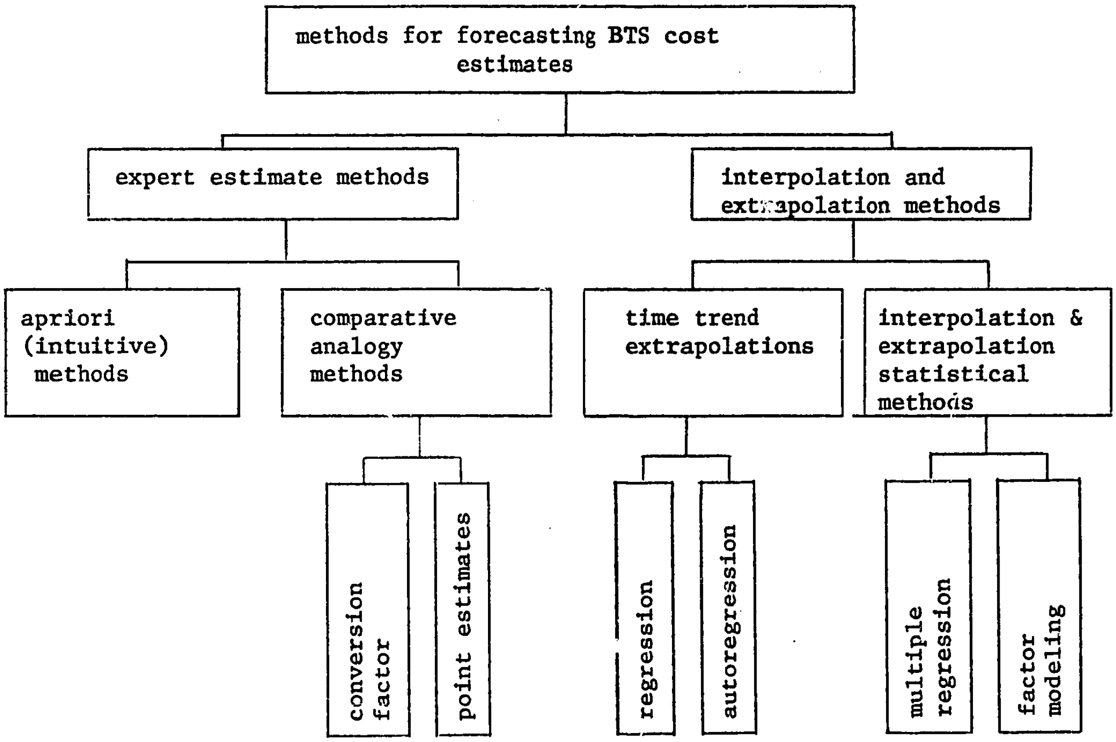


Fig. 4.6. Classification of methods for forecasting the cost estimates depending upon sources of forecast information

However, this does not disrupt the strictness of classification, since the role of the expert is restricted to incorporating his corrections in that information which the model produces. Thus, here it is appropriate to speak not about forecasting as such but rather about increasing the reliability of the interpolation and extrapolation forecast estimates.

In the general instance the choice of the forecasting method for the value estimates is dictated by the considerations of forecast reliability and dependability. From this viewpoint, the priority of the methods employed as a forecasting tool for the model of the cost estimates is beyond any doubt. However, their use is not always possible, since in order to obtain the model there must be sufficiently representative information characterizing the development processes of the cost estimate's object. As was shown in Chapter 1 (see Fig. 1.4), any functional element in its development passes through a series of sequential phases: the phase of embodiment which precedes the appearance of the prototype; the initial phase of development encompassing the period from the appearance of the prototype to the first example used in a functioning system; the phase of intensive development characterized by rapid growth of the element's functional parameters and its penetration into ever-newer systems; the obsolescence phase characterized by a sharp drop in the growth rate of the functional parameters and this ultimately leads to the halting of scientific research and development in the given area. Simultaneously with the object's development there is the accumulation of information on the articles made or the prototypes of the forecast object. This information includes data on the functional characteristics of the prototypes, the design solutions corresponding to these characteristics, the production processes, forms and methods of organizing development and production, external factors and so forth.

The consolidated characteristics of the initial data about the cost estimate's object and the contents of the typical value forecasts in comparison with the forecast methods which are most characteristic for the listed development phases of system functional elements are shown in Table 4.1. As is seen from the table, it is necessary to resort to expert judgments based largely on expert intuition during the phase preceding the appearance of the prototype, that is, when the general operating principles of the system's element are known. Such a situation could arise, for example, prior to the appearance of the first liquid-fuel rocket engine or the first computer.

At this stage, only the principle for creating the thrust (for the liquid rocket engine) or transforming information (for the computer) is known but in essence there is not a single extant full-scale model of the new functional element. For this reason the information on the new object is very meager as largely it involved results of theoretical and experimental research. In such an instance it is extremely difficult to forecast the technical possibilities of the new device let alone a cost forecast.

Either type of forecast can be made only by experts and the soundness of such experts will be very low. However, regardless of the probable occurrence of major errors in the forecast estimates, under such circumstances no other alternative simply exists. It is better to have a bad forecast than none at all.

The appearance of the prototypes of new functional elements provides the researcher with enough information to markedly increase the forecast accuracy. At this stage the basic functional parameters of the article are already known as well as the particular features of design and the basic types of materials and production processes. The actual expenditures are very important information from the standpoint of the cost forecast and would include the cost of the scientific research and development carried out before the construction of the prototype, the cost of the prototype, the cost structure and so forth.

Under these conditions, the forecast of the cost estimates for a system element possessing new, higher functional characteristics becomes more dependable. However, it must be remembered that the first examples of new functional elements are only the prototype of future articles, that is, the articles which will go into serial production and operation and for this reason the ambituity of the forecast estimates is still great.

Nevertheless, formalized methods can be employed for forecasting the cost estimates. These are methods based on analogies or, as they are sometimes called, comparative forecasting methods. The essence of these methods is that the production cost of the nearest predecessor (prototype) of the cost forecast object is employed as the initial base for forecasting. However the quantitative estimate of the indicators which make up the forecasting tool is carried out by experts.

Table 4.1

development phases of BTS functional element [1]	content of initial statistical data on object development processes [2]	content of forecast information on cost estimates [3]	forecasting methods [4]	forecast information sources [5]
Phase preceding appearance of prototype (embodiment phase)	1. Physicochemical functioning principles, certain methods of creating useful effect based on results of theoretical & experimental research, etc.	1. Cost of NIOKR for basically new functional element	Intuitive, expert estimates	Experts
	2. Possible spheres of use--type of systems in which new functional element can be used	2. Cost of developing serial production and operation of new functional element	same	same
	3. General ideas on effect from using new functional element in operating (existing) systems	3. Cost of serial production and operation of new functional element	same	same
Initial development phase--from appearance of 1st functioning prototype to 1st industrial models	1. Basic functional parameters, design features, basic production processes, materials	1. Cost of NIOKR carried out for stabilizing functional properties of elements	Expert estimates based on analogues	same
	2. Actual cost of scientific research and development carried out before creation of prototype & its structure	2. Cost of NIOKR of systems using new functional element		
	3. Costs of manufacturing prototype (prototypes) and its structure			

Table 4.1 (Con't.)

[1]	[2]	[3]	[4]	[5]
Intensive development phase from 1st industrial models to start of saturation period	1. Characteristics of certain aggregate of objects which have undergone all life cycle stages: NIOKR, serial production (SP) and operation	1. Cost of NIOKR, serial production and operation of system using alternatives of functional element	Time & statistical extrapolation	Models
	2. State characteristics of life cycle stages for years of object's development	2. Modification costs	Statistical extrapolation, comparative methods	Models, experts
	3. Design efficiency characteristics of made articles	3. Cost of NIOKR, SP & operations of various design solutions for functional element	Conversion factor method, point estimate method, direct calculation methods	Models, experts, standards
	4. Actual expenditures, their structure by expenditure items and economic elements			
Obsolescence stage from start of saturation to full halt of NIOKR in given area	1. Same as in preceding stage but on larger scale	1. Cost of NIOKR to broaden theoretical limits of functional element	Expert estimates	Experts
	2. Theoretical limits to change in functional characteristics	2. Cost of NIOKR for coming closer to limit values of functional characteristics 3. Points 1, 2, 3 of preceding phase	Time and statistical extrapolation combined with expert estimates	Models, experts

As information is acquired on the object's development processes and the formation of the cost estimates, an opportunity arises for establishing the regular changes in the cost estimates under the influence of scientific and technical progress. A mathematical model is constructed from these data and its describes the process of forming the cost estimates and the interpolation and extrapolation calculations are made depending upon the forecasting goals.

Interpolation and extrapolation forecasting methods can be employed in using two types of models: models for the time dynamics of the resulting trend of the cost estimates (see Fig. 4.5) and regression models which reflect the basic relations between the cost estimates and the factors of their formation.

In the first instance it is assumed that the dynamics of the cost estimates are a consequence of a self-developing process engendered by factors which are internally inherent to the processes of scientific and technical development. The interaction of the factors forming the resulting trend is considered to be a reaction to the necessity of ensuring the required functional parameters with the least expenditures of social labor. For this reason the existing time trend for the cost estimates can be a forecast tool. In the second instance an attempt is made to describe the process of forming the cost estimates using a model which would reflect as fully as possible the relations of the cost estimates with the factors of each group. Here it is assumed that quantitative relationships exist between the cost estimates and the corresponding factors. In either instance the concept of the inertia in the processes of forming the cost estimates is used. This means that the relations observed in the past, that is, over the object's developmental prehistory, will remain in the future.

Cost estimates can be forecast with a varying degree of detailing for calculations. Depending upon the degree of detailing, forecasting methods are divided into aggregated and differentiated (Fig. 4.7). The detailing of the calculations can aim at two goals: increasing the accuracy of the forecast estimates or determining expenditures according to the type of consumed resources.

In the first instance the detailing is carried out in two areas: 1) the articulation of the object into individual design elements (units, assemblies or parts) and 2) the structurizing of the individual life cycle stages. It is assumed that detailing in terms of design elements provides an opportunity to consider the influence of parameters which determine consumer value of the assemblies, units and parts on their cost estimates (for example, the increased pressure in a compressor or gas temperature ahead of a jet engine turbine). Structurization of the life cycle stages makes it possible to reflect the influence of specific features of individual events in the life cycle on the cost estimates (for example, the breaking down of the serial production process into three stages: preparation, starting up and full serial production).

The forecasting of expenditures by resource types is carried out broken down for the economic elements comprising the cost structure of the article. Here the cost of the material, labor and other types of expenditures is determined. Each of the designated varieties of the differentiated methods, in turn, is divided into two types: methods for direct integration of the estimates and methods for a base position or proportional ratios.

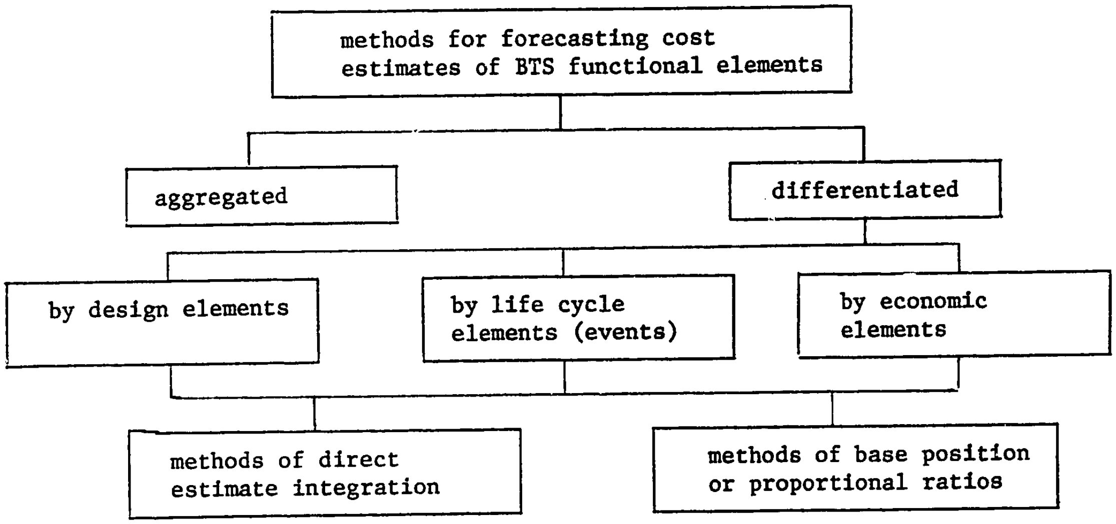


Fig. 4.7. Classification of forecasting methods for cost estimates depending upon degree of calculation detail

The method of direct integration of the estimates envisages the direct totaling of the cost estimates of each element in accord with the adopted structure of the functional element, the life cycle stage or cost. The use of the base position methods makes it possible to restrict oneself to forecasting the cost estimates for the basic structural elements (for example, the basic assembly, the cost of basic materials, the production costs of a prototype) and so forth. The cost estimates of all remaining elements are determined from the ratio to the cost estimate of the base position using the standard expenditure structures. The base position method is less accurate in comparison with the direct integrating methods as the cost estimate structures are subject to significant changes. However, in the absence of proper accounting of expenditures for all the elements, this method can be useful, although for very approximate estimates.

Generally speaking the possibility of employing differentiated methods depends largely upon the system for calculating the actual expenditures on the creation, production and operation of the systems. Thus, if expenditures on manufacturing a functional element are calculated as a whole for the article, then forecasting the cost estimates in terms of design elements is excluded. Moreover, the use of differentiated methods entails the carrying out of cumbersome calculations and this is not always justified by higher estimate accuracy. For this reason the aggregated cost estimate forecasting methods have become most widely used while the differentiated calculations are widely used in planning.

From what has been said it follows that the choice of the forecasting method for the cost estimates depends chiefly upon the scope and composition of initial information about the object of the cost estimate. The reliability of the forecast estimates depends largely on this.

However, it is essential to point out that the selection of the forecasting method for the cost estimates can also be determined by the goals of the forecast research. Thus, in solving the problems of choosing optimum system parameters and other characteristics of development processes, the forecast model should reflect the link of the cost estimates with each of the optimized characteristics. The constructing of such a model requires the use of multiple regression and factor modeling methods and this in and of itself, as we will see subsequently, involves the need to solve a whole series of complex conceptual problems and carry out labor intensive and cumbersome calculations. For this reason, in those instances when the influence of an individual characteristic on the cost estimates is not major, for example, in setting prices, determining the overall need for R&D financing and so forth, preference can be given to simpler methods which, in particular, would include the time trend extrapolation methods.

Let us move on to a more detailed examination of the methods for forecasting the cost estimates of BTS functional elements, paying basic attention to the analogy and factor modeling methods. All the remaining methods of the designated classification (see Fig. 4.6), from the viewpoint of the formal forecasting apparatus, in no way differ from the methods which were described in Chapter 2.

4.2.2. Comparative Forecasting Methods of Cost Estimates

The comparative methods or the analogy methods are based upon a comparison of the parameters for the object of the cost estimate with the parameters of its prototype (hence the name of the methods). Of greatest interest are two varieties of the comparative methods: the point estimate method and the conversion factor method.

The essence of the point estimate method is that a unit of increase for a certain functional parameter P_{f_i} is given a cost number C_{b_i}. Here the forecasting of the cost estimate for the object is carried out using the formula

$$C = C_p + \sum_{i=1}^{n} C_{b_i} P_{f_i}, \qquad (4.3)$$

where C_p--prototype production costs.

The cost number indicates by how much the production cost of a prototype can increase if the functional parameter increases by one. The model of the point estimate is formed using experts who assign a number to each parameter (for example, according to a 100-point scale):

$$b_1 \to P_{f_1}, \quad b_2 \to P_{f_2}, \quad \ldots, \quad b_n \to P_{f_n}. \qquad (4.4)$$

The transition from the numbers to the cost number in the model (4.3) is carried out by norming the point estimates

$$C_{b_i} = b_i \Big/ \sum_{i=1}^{n} b_i. \qquad (4.5)$$

The basic drawback of the method is that absolute parameter increases are used in the model (4.3). As a consequence of this their dimensionality influences the amount of the production costs. Here greater preference will be given to the parameters measured in absolute units in comparison with parameters measured in relative units and this leads to significant distortions of the forecast estimates.

The most objective of the comparative methods is the conversion factor method. The production cost of an object here is determined as the product of the prototype production costs by the average conversion factor:

$$C = C_p K_{p_f}. \tag{4.6}$$

The mean conversion factor K_{pf} is determined by the particular conversion factors from the prototype parameters to the parameters of the forecast object as weighted in accord with the degree to which the parameter influences the amount of the cost estimate.

The mean conversion factors can be expressed by linear and nonlinear models. Here the model variables are represented in the form of either indexes or specific increments. A linear model for the mean conversion factors with variables having the form of indexes is determined by the expression

$$K_{p_f} = \sum_{i=0}^{n} \beta_i K_{p_{f_i}}, \tag{4.7}$$

where β_i--the weight factor considering the proportional (specific) influence of the parameter in forming the cost estimate;

K_{pf_i}--the particular conversion factor representing the ratio of the estimated object's parameter to the prototype parameter:

$$K_{p_{f_i}} = P''_{f_i} : P'_{f_i}. \tag{4.8}$$

A model employing proportional parameter increments is written thus:

$$K_{p_f} = 1 + \sum_{i=1}^{n} \beta_i \frac{\Delta P_{f_i}}{P'_{f_i}}. \tag{4.9}$$

The weight factors which consider the influence of the parameters of the forecast object on the cost estimates are determined by expert estimates. The total of the weight factors in (4.7) and (4.9) should satisfy the condition $\sum\limits_i \beta_i = 1$, as the overall influence of the parameters on expenditures is considered as one while the contribution of each parameter to forming the cost estimates is measured in fractions of a unit. In this regard in forming the model (4.7), there should be the certainty that without exception all the parameters have been considered which influence the cost estimates. At the same time, in analyzing the model (4.9), it can easily be noted that for the parameters close in size the product $\beta_i(\Delta P_{f_i}/P_{f_i}) \approx 0$.

This frees one from the need to examine the entire aggregate of parameters, since the results of the forecast will be influenced only by the parameters which have undergone a marked change. As a consequence, the number of parameters which must be studied is sharply reduced and this can substantially simplify the expert estimate procedure.

Thus, it can be stated that the proportional increment model is a particular case of the model (4.7). However, its use in the form it is shown in (4.8) does not fully eliminate the problem of narrowing the number of considered variables. For this reason an important step which should precede the expert estimates is the isolating of the base parameter which determines the overall level of expenditures on the given type and purpose of functional element. Among such base parameters are, for example, the weight of the element or other characteristics more closely linked to its specific purpose (the weight of an aircraft, engine thrust, pulse power of a radar and so forth).

The isolating of the base parameter makes it possible to move on to forecasting proportional production costs, that is, the production costs of a unit of the base parameter Cp_0. This transition in and of itself provides an important advantage, as for a narrow class of articles the base parameter may not change substantially and technical progress will be determined by the dynamics of qualitative characteristics (for example, aircraft speed, specific engine fuel consumption, the sensitivity of radio equipment and so forth).

The advantages of models using base parameters become particularly apparent if the following model is used for analysis:

$$C = P_0'' C'_{P_0} \left(1 + \sum_{i=1}^{n} \beta_i \frac{\Delta P_{f_i}}{P'_{f_i}} \right), \qquad (4.10)$$

where P_0''--the amount of the base parameter for the cost estimate's object;
C'_{P_0}--the proportional cost of a unit of prototype base parameter.

In making the transformation, we obtain

$$C = P_0'' C'_{P_0} + P_0'' C'_{P_0} \sum_{i=1}^{n} \beta_i \frac{\Delta P_{f_i}}{P'_{f_i}}. \qquad (4.11)$$

From expression (4.11) it is seen that the object's production costs are made up of two components. The first part is the production costs which depend upon the amount of the base parameter P_0'' and considers the amount of all the remaining parameters which have played their role in forming the prototype production costs (the latter is expressed in the amount C'_{P_0}). The second part considers the change in prototype production costs as a consequence of the dynamism of quality performance. From this it follows that mistakes in setting the weights are reflected only in a part of the cost estimate and namely the one which in the righthand part of (4.11) is expressed by the second number of the total. It is natural to assume that with other conditions being equal the accuracy of the models with the isolated base parameter will be higher.

The designated models of conversion factors are based on the assumption that the relation between the cost estimates and the characteristics determining them is close to linear. At the same time, research has shown that the influence of a majority of factors considered by conversion factors on the cost estimates is nonlinear and for this reason linear models are better used with relatively small changes in the variables. Otherwise nonlinear ones must be used, for example, the multiplicative models of the conversion factors:

$$K_{P_l} = \prod_{i=1}^{n} K_{P_{l_i}}^{\beta_i}. \qquad (4.12)$$

The model (4.12) is linear but relative to the logs of the variables. Having logged both parts of the equation, we obtain

$$\log K_{P_l} = \sum_{i=1}^{n} \beta_i \log K_{P_{l_i}}. \qquad (4.13)$$

As is seen from (4.13), the particular conversion factors are represented by the logs of the parameter ratios and their significance has the sense of the weight factors of the ratio logs. In this instance, it is essential to maintain the rule $\sum_{i=1}^{n} {}_i = 1$. As a whole it must be pointed out that the conversion factor method is enticing in its simplicity and convincingness. However its successful application depends largely upon the skill in setting the weights which reflect the contribution of one or another characteristic in the process of forming the cost estimates. The total of the weights reflecting the influence of the designated factors on the cost estimates should be equal to one. Under these conditions the problem comes down to distributing the general influence of the factors taken as one among the individual factors. Formally it must be established how much more strongly one factor influences the cost estimate in comparison with the others, that is, bring out the priority relationships between the factors.

Of the presently known methods for establishing priorities, it is essential to mention the methods of successive preferences [6] and paired comparisons [45] which can be successfully applied for establishing weights in the conversion factor models. The choice of one of them, as practice shows, can be made depending upon the number of factors to be compared. With a rise in the number of factors to be compared p, the mentioned methods can be broken down into the following sequence: successive preferences ($p \leqslant 6$); paired comparisons ($p > 6$). Let us take up the essence of these methods.

The successive preferences method provides a successive comparison of the expert estimates considering the following assumptions: for each factor there is a corresponding real nonnegative number γ_j viewed as the estimate for the true significance of the factor 0_j (subsequently for simplicity of exposition by the term the 0_j factor we will understand its true significance); if the 0_j factor is more important than 0_k, then $\gamma_j > \gamma_k$, and if 0_j equals 0_k, then $\gamma_j = \gamma_k$. The procedure of successive comparisons consists of three stages.

1. The analyzed factors are ranked, that is, placed in the order of diminishing preferences, and here rank 1 receives the most preferred factor and rank p the least preferred. The most important factor is assigned a weight $\gamma_1 = 1$, while the remaining factors receive estimates γ_i between 0 and 1 in the order of their relative importance.

2. In the second stage the following is established: a) Is factor 0_1 more important than all the remaining factors, that is, a comparison is made of 0_1 with $0_2 + 0_3 + \ldots + 0_p$:

If 0_1 is preferable to $0_2 + 0_3 + \ldots + 0_p$, it is essential to alter the significance of γ_1 so that an inequality is met: $\gamma_1 > \gamma_2 + \gamma_3 + \ldots + \gamma_p$;

If 0_1 and $0_2 + 0_3 + \ldots + 0_p$ are equal, then the significance of γ_1 must be altered so as to meet the equality $\gamma_1 = \gamma_2 + \gamma_3 + \ldots + \gamma_p$;

If 0_1 is less preferable than $0_2 + 0_3 + \ldots + 0_p$, then a value of γ_1 is adopted so as to meet the inequality $\gamma_1 < \gamma_2 + \gamma_3 + \ldots + \gamma_p$;

b) Is the second in importance factor 0_2 with an estimate γ_2 more preferable than all the remaining factors which have received lower estimates, that is, 0_2 is compared with $0_3 + 0_4 + \ldots + 0_p$. Depending upon the results of the comparison, the significance of γ_2 is corrected using the same rules as for γ_1. The procedure of successive comparisons is continued up to the (p - 1) factor.

After this the entire procedure is repeated as the quantitative estimates obtained in the first cycle for the significance of the factors cannot satisfy the initially set priorities $O_1 \gtrless \sum_{i=2}^{p-1} O_i$, $O_2 \gtrless \sum_{i=3}^{p-2} O_i$ and so forth. Thus it may seem that with $O_1 > \sum_{i=2}^{p-1} O_i$ $\gamma_1 \leqslant \sum_{i=2}^{p-1} \gamma_i$ or with $O_1 < \sum_{i=2}^{p-1} O_i \gamma_1 \geqslant \sum_{i=2}^{p-1} \gamma_i$ and so forth. In this regard, there must be several iterations for matching the quantitative estimates for factor significance and the established priorities and this extends the calculation procedure of the weight factors.

For simplifying the successive preference procedure, the comparison procedure can be altered so that the comparison and reestimating of the factors are made in the reverse order: initially there is a comparison of the factor 0_{p-2} with $0_{p-1} + 0_p$, then the factor 0_{p-3} with $0_{p-2} + 0_{p-1} + 0_p$ and so forth. The process is continued until a comparison has been made of the factor 0_1 with $0_2 + 0_3 + \ldots + 0_p$. The use of such a method makes it possible not to check the obtained results and this significantly simplifies the successive comparisons procedure.

3. The obtained preference estimates are normed. The normed preference estimates can be used as weights in the conversion factor models.

The paired comparisons method consists in a paired comparison between two studied factors so as to establish the more important factor in each pair. For facilitating this procedure, a paired comparisons matrix is drawn up and here all the factors

(1, 2, ..., p) are written twice in the same order: in the top line and in the extreme lefthand column as is shown in Table 4.2.

Table 4.2

factors	factors					
	1	2	...	j	...	p
1	▥	x_{12}	...	x_{1j}	...	x_{1p}
2	x_{21}	▥	...	x_{2j}	...	x_{2p}
...	...	...	...		...	
i	x_{i1}	x_{i2}	...	x_{ij}	...	x_{ip}
...	...	...	...	...	...	...
p	x_{p1}	x_{p2}	...	x_{pj}	...	▥

Each expert who fills out such a matrix should put an estimate x_{ij} at the intersection of the line and column for the two compared factors. Depending upon whether factor i is more preferential than factor j, this estimate equals 1 or 0, respectively.

After each expert has filled out such a matrix, a second matrix is constructed showing a percentage relationship of the instances when factor i was more important than factor j in the total number of obtained estimates (Table 4.3).

The elements of the matrix given in Table 4.3 possess the property that $\lambda_{ij} = m_{ij}:m$, where m_{ij}--the number of experts who preferred factor i in comparison with factor j; m--the number of experts participating in the expert questioning. Thus, if the number of experts m = 10, and four experts preferred factor i, then for $\lambda_{ij} = 4:10$, and, respectively, six experts preferred factor j, then $\lambda_{ji} = 6:10$, hence $\lambda_{ij} + \lambda_{ji} = 1$.

After obtaining the generalized preference matrix the elements of which λ_{ij} represent the relative number of preferences for each pair of factors obtained from all the experts, the estimates are then scaled. The procedure for constructing the scale estimates consists in turning the observed ratios λ_{ij} into the most probable ones z_{ij}. This transformation is made according to the formula

$$G(z_{ij}) = \lambda_{ij} = \int_{-\infty}^{z_{ij}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \qquad (4.14)$$

where z_{ij}--the normed deviation of the observed share of instances of preferring factor i to factor j from the true share [8].

Table 4.3

factors	factors						
	1	2	...	j	...	p	series total
1	▨	λ_{12}	...	λ_{1j}	...	λ_{1p}	λ_1
2	λ_{21}	▨	...	λ_{2j}	...	λ_{2p}	λ_2
...	...	...	...	...	...	...	...
i	λ_{i1}	λ_{i2}	...	λ_{ij}	...	λ_{ip}	λ_i
...	...	...	...	...	...	...	...
p	λ_{p1}	λ_{p2}	...	λ_{pj}	...	▨	λ_p

Here λ_{ij} is viewed as the area of the normed normal distribution from $-\infty$ to z_{ij}. The values of z_{ij} comprise the basic transformation matrix with rows of numbers for each factor i and columns of numbers for each factor j as is shown in Table 4.4. In the given matrix, each estimate of z_{ij} is the difference between the factor i and the factor j in the standard deviations and the total of these estimates

$z_i = \sum_{j=1}^{p} z_{ij}$, while the mean value is

$$\bar{z}_i = \sum_{j=1}^{p} z_{ij}/m, \qquad (4.15)$$

where p--the number of analyzed factor.

For each factor, from the mean value of $\bar{z}_i$, using the table for the normed normal distribution in [8], $G(\bar{z}_i)$ is determined and on this basis the normed relative importance coefficients as well. Here the estimate $G(\bar{z}_i)$ are totaled for all factors and each estimate is divided by the obtained total:

$$\beta_i = G(\bar{z}_i) \Big/ \sum_{i=1}^{p} G(\bar{z}_i); \quad \sum_{i=1}^{p} \beta_i = 1. \qquad (4.16)$$

The amounts necessary for calculating the relative importance coefficients can be reduced to Table 4.5.

Table 4.4

factors	factors							
	1	2	...	j	...	p	total	average
1	▨	z_{12}	...	z_{1j}	...	z_{1p}	z_1	$\bar{z}_1$
2	z_{21}	▨	...	z_{2j}	...	z_{2p}	z_2	$\bar{z}_2$
...	...	...	...	...	...	...	...	...
i	z_{i1}	z_{i2}	...	z_{ij}	...	z_{ip}	z_i	$\bar{z}_i$
...	...	...	...	...	...	...	...	...
p	z_{p1}	z_{p2}	...	z_{pj}	...	▨	z_p	$\bar{z}_p$

Table 4.5

factors	$\bar{z}$	$G(\bar{z})$	normed relative importance coeffs.
1	$\bar{z}_1$	$G(\bar{z}_1)$	β_1
2	$\bar{z}_2$	$G(\bar{z}_2)$	β_2
...	...	...	...
i	$\bar{z}_i$	$G(\bar{z}_i)$	β_i
...	...	...	...
p	$\bar{z}_p$	$G(\bar{z}_p)$	β_p

The paired comparisons method, when the priorities are set using the estimates 1 and 0 assigned to the analyzed factors, requires the deviation significant estimates in the opinions of the experts relative to the preferences set by them. Considering that the significance estimate presupposes a normal deviation distribution, such a procedure requires that an expert group which is sufficiently representational in size be involved in the expert evaluation.

According to certain data [6, 36], the number of experts should be at least 25. This condition imposes a constraint on the possibility of using the paired comparisons method with the assigning of the extreme preference estimates (1, 0). For this reason with an insufficient size of the expert groups it is possible to propose that the estimating of the weights in the conversion factor models be made using preferences set according to a 10-point scale.

In this instance each expert fills out a matrix analogous to the one shown in Table 4.6, only instead of the estimates 1 and 0, point estimates are set in the table squares. Here, if a factor standing in line i, in comparison with the factor standing in column j is considered more significant, then it is given an estimate $y_{ij} > 5$. Then the factor standing in column j is given the estimate $y_{ji} = 10 - y_{ij}$. Here the number of points are set in whole numbers. After carrying out the expert evaluation, the table's data are generalized. For this the total number of points is determined as set by expert k for each line i $\sum_{i=1}^{p} y_{ijk}$; the total number of points set by expert k for each column j $\sum_{i=1}^{p} y_{ijk}$; the resulting total of points $\sum_{i=1}^{p} \sum_{j=1}^{p} y_{ijk}$.

The weight coefficients set by expert k for each factor i are determined from the following formula

$$\beta_{ik} = \sum_{j=1}^{p} y_{ijk} \Big| \sum_{i=1}^{p} \sum_{j=1}^{p} y_{ijk}. \tag{4.17}$$

The mean statistical value of the weight coefficient, in being the indicator of the generalized opinion of the experts, is determined from the formula

$$\beta_i = \sum_{k=1}^{m} \beta_{ik}/m. \tag{4.18}$$

As we see, the procedure for calculating the weight factors using a 10-point scale is greatly simplified. At the same time, the use of this method requires additional procedures for establishing the reliability of the expert estimates as was pointed out in Chapter 2.

The elucidating of the factor priorities from the standpoint of their impact on the cost estimates using the just described methods, in addition to the procedural difficulties, entails also a number of logical and psychological difficulties. The latter are determined primarily by the absence of cause-and-effect relations between the factors and the cost estimates.

As was shown in Section 4.1, the operational effectiveness factors influence the cost estimates of the system functional elements not directly but rather indirectly through a multiplicity of intermediate characteristics. Thus, the operational effectiveness factors influence the design characteristics of functional elements and the physical properties of the employed materials. The latter determine the state of the production processes, the means and implements of labor and this, in turn,

Table 4.6

factors	factors							
	1	2	...	j	...	p	total points	weight coeff. for **i** factor
1	▨	y_{12}	...	y_{1j}	...	y_{1p}	$\sum_{j=1}^{p} y_{1j}$	β_{1j}
2	y_{21}	▨	...	y_{2j}	...	y_{2p}	$\sum_{j=1}^{p} y_{2j}$	β_{2j}
...	...	...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ip}	$\sum_{j=1}^{p} y_{ij}$	β_{ij}
...	...	...	...	...	...	...	...	...
p	y_{p1}	y_{p2}	...	y_{pj}	...	▨	$\sum_{j=1}^{p} y_{pj}$	β_{pj}
total points	$\sum_{i=1}^{p} y_{i1}$	$\sum_{i=1}^{p} y_{i2}$	...	$\sum_{i=1}^{p} y_{ij}$	...	$\sum_{i=1}^{p} y_{ip}$	$\sum_{i=1}^{p}\sum_{j=1}^{p} y_{ij}$	

influences the labor intensiveness of the work, the skill level of the workers employed in these jobs, the cost of the employed means and implements of labor, material intensiveness and so forth. Only the last group of characteristics has a direct influence on the amount of the cost estimates. Under these conditions the attempt with expert help to set the influence of one or another functional characteristic of a system's element on its cost estimates is directly tied to the necessity of overcoming the ambiguities concerning the multiplicity of intermediate links, their interaction and reciprocal influence. The following circumstance must also be pointed out. Due to the hierarchical nature of the functional structure of technical systems, the design characteristics and the properties of the employed materials in the functional elements are determined not only by the parameters of the elements themselves but also by the parameters of the higher level element in which the given system is a subsystem.

In addition, a number of design and materials characteristics can be determined by the influence of functional parameters in the elements of the same hierarchical level as the cost estimate's object. This necessitates the creating of an information basis and the elaborating of special procedures for an expert estimate of the weights in the conversion factor models. Obviously the information basis which precedes the expert evaluation should be an hierarchical system of interrelated characteristics which reflects the mechanism of influence of the employed factors on the cost estimates through all the intermediate characteristics. In having such a system it is possible to organize a step-by-step expert evaluation, in estimating in sequence the priorities between the factors which are on the given level in relation to the characteristics of neighboring levels.

Here for the estimates on each level or several adjacent levels for which the estimates involve the solving of kindred problems, independent expert groups can be set up. Each such group will consist of experts from one or two related specialties and this will make it possible to substantially increase the reliability of the expert estimates. The estimate reliability can also be increased by correctly choosing the method of disclosing the priorities since the number of estimated factors, like the number of experts in each individual instance, will not be constant.

The hierarchical system reflecting the mechanism of influence of the operational effectiveness factors on the cost estimates and incorporating the main intermediate links between the characteristics of the various cosubordination levels has been worked out by the authors with the participation of the engineers Yu. A. Teplov and V. I. Ul'yanov. This system called an influence graph for the operational effectiveness factors on the cost estimates is shown in Fig. 4.8.

As is seen from the diagram, on the first and second levels of the hierarchy there are characteristics of the technical system in which the assessed functional element is a subsystem. Here it is shown that the design characteristics of a functional element and the properties of the employed materials are determined not only by the technical, operational and functional characteristics of the element but also by the system's analogous characteristics. Then the diagram shows the successive influence of the design characteristics of the functional elements and the properties of the employed materials on the complexity characteristics of the production processes (for example, the method of obtaining stock, the machining method, the method of connecting the elements and so forth). The designated characteristics of the production processes, in turn, are related to the basic expenditures which determine the production cost of a unit of the base parameter for the functional element.

The procedure of expert estimates made using the influence graph should consist of disclosing the specific influence of the upper level characteristics on the inferior level characteristics for all interrelated groups. Here the general influence of one or several upper level characteristic groups on each group of the related lower level characteristics should equal one. It is not hard to see that all the characteristics are interrelated by cause-and-effect ties and this sharply reduces the degree of ambiguity on the intercausality of the factors.

In addition, the link between any two groups of factors can be established by an isolated independent expert group. The latter provides an opportunity to set up the necessary number of uniform expert groups, each of which will consist of the representatives of related specialties. Thus, the relationship between the functional characteristics of the system and the element can be set by specialists in systems theory; the relations between the functional and design characteristics and the physical properties of the employed materials will be set by designers; the relation between the physical and production properties of the materials by production engineers and so forth.

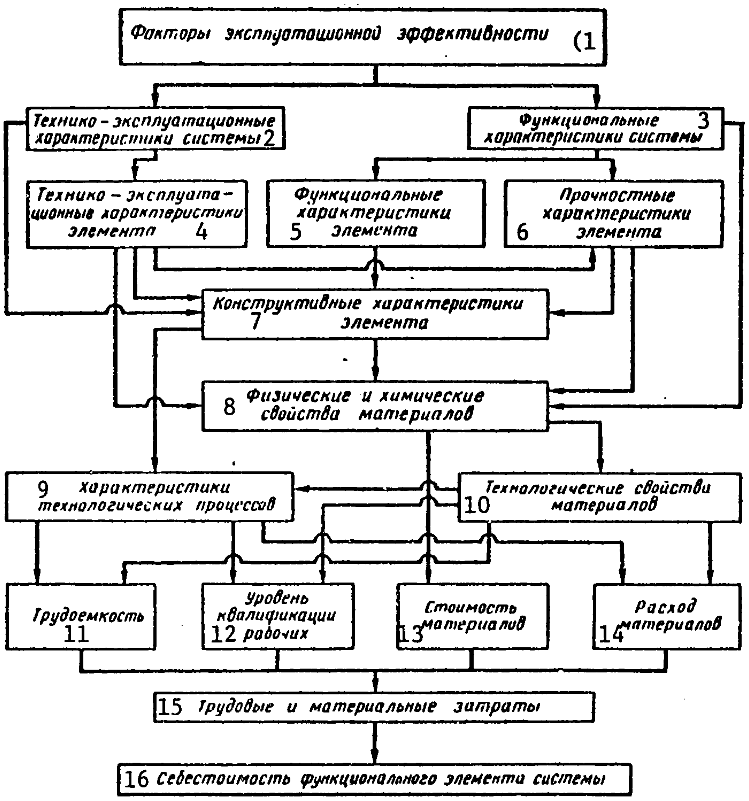


Fig. 4.8. Influence graph of operational effectiveness factors on cost estimates

Key: 1--Operational effectiveness factor; 2--System's technical and operating characteristics; 3--System's functional characteristics; 4--Technical and operational characteristics of element; 5--Functional characteristics of element; 6--Strength characteristics of element; 7--Design characteristics of element; 8--Physical and chemical properties of materials; 9--Characteristics of production processes; 10--Technological properties of materials; 11--Labor intensiveness; 12--Worker skill level; 13--Cost of materials; 14--Consumption of materials; 15--Labor and material expenditures; 16--Production cost of a functional element for system

The organizational difficulties of conducting such an expert evaluation are apparent, however there obviously is no other reasonable alternative for constructing the conversion factor models. At the same time the conversion factor models are a very valuable and at times irreplacable economic forecasting tool since, as a rule, the number of the characteristics of the technical systems influencing their cost estimates is great while the number of prototypes comprising the initial statistical aggregate in the majority of instances is insignificant. This circumstance imposes constraints on the possibility of constructing the statistical dependences which will be taken up in the following section.

4.2.3. Interpolation and Extrapolation Methods for Statistical Dependences

The methods of interpolation and extrapolation of statistical dependences are based on the assumption that there are quantitative relations between the value estimates of the BTS functional elements and their characteristics as well as the basic factors describing the state of the individual stages of the BTS life cycle and the environment. The task usually is to disclose these ties, to select the main factors out of the multiplicity of them, to localize the ties which cannot be estimated quantitatively and to coordinate the entire aggregate of determining factors by a mathematical model which would reflect the basic patterns of the studied phenomenon. The models which satisfy these requirements have been named the mathematical economics cost models and at present are the basic forecasting tool for the cost estimates.

The essence of forecasting using the mathematical economics cost models is that by the statistical dependences which reflect the influence of the indicators of interest to us on the cost estimate, the probable changes of the cost estimates are determined when the indicators assume new values different from those which were observed in the initial statistical aggregate. Here, if the new value of the indicator does not go beyond the limits of the observed range of changes, the forecast is of an interpolation nature. Otherwise the set statistical dependence between the cost estimate and the changeable indicator is extrapolated for its new values.

In the BTS economic forecasting problems, of greatest interest is the extrapolation of statistical dependences since the characteristics of developmental processes evolve. However in the problems of optimizing the BTS parameters, the need arises also for interpolation calculations, if the optimization process is carried out under a certain compromise scheme. In this instance, the optimal variation of a system can include functional elements the individual parameters of which will be somewhat below the parameters of the nearest prototypes. In order to assess the expenditures on such functional elements, an interpolation of the statistical dependences is employed.

The mathematical economics cost models can be of two types. The model can be represented by one general multiple regression equation reflecting the influence of the entire aggregate of factors simultaneously. This is the simplest method of modeling the cost estimates and it provides very rough forecasts. More dependable forecasts are achieved by employing composite, mixed models based upon factor modeling of the cost estimates. The factor modeling method consists in modeling the influence of the individual groups of related factors on the cost estimates separately in accord with the place, time and nature of this influence. The thereby obtained local models or submodels are brought together according to definite rules into the overall mathematical economics model.

The use of factor modeling methods becomes possible with a sufficiently full and representative amount of statistical information encompassing the various aspects in the developmental process of the cost estimate's object over large intervals of time. Since these conditions are not always feasible, at times one must resort to multiple regression equations.

At the same time it is essential to point out that regardless of what the final form is for showing the mathematical economics model of the cost estimates or the form of the multiple regression equation or composite model, the modeling of the cost estimates must start by dividing the general process of forming the cost estimates into individual fragments, the elucidation of the composition of influencing factors and the establishing of quantitative indicators reflecting the influence of the factors on the cost estimates.

For this reason the constructing of one or another mathematical economics model for cost estimates must be carried out according to a general scheme (Fig. 4.9). The entire sequence of events represented in the basic block diagram for modeling the cost estimates can be divided into a series of independent stages: logical modeling (blocks 1-2), the forming of the initial information file (block 3), logical-statistical analysis or selection of influencing factors (blocks 4-5), mathematical modeling (block 6), forecasting (block 7) and forecast verification (block 8).

The logical modeling stage envisages the carrying out of preliminary research aimed at showing the possible states of the cost estimate's object, the stages of its life cycle and of the environment. In this stage a logical scheme is constructed making it possible to compile a general notion about the formative mechanism of the cost estimates under the influence of the developmental processes.

The logical modeling of the states of the cost estimate's object (block 1.1) is carried out in the following sequence: the class of systems is established which are to include the cost estimate's object; the system prototypes are selected which are identical in terms of their basic purpose and performed functions; the system is decomposed with a model being constructed for the internal structure and the place and role of the object in the process of the system's functioning are determined; the parametric series of the object's prototypes is set; the basic functional characteristics of the object, their internal and external relations are set; the object is decomposed with its schematic diagram, composition and purpose of the basic structural elements and other design characteristics being determined; the relations are established between the basic functional and design characteristics; the object's model is constructed showing its structure, internal and external relations.

In analyzing the functional characteristics, particular attention must be paid to their ties with the elements of the same and adjacent hierarchical levels of the system. This is essential for fully encompassing the factors which influence the development of the cost estimate's object, as the latter is often determined precisely by the external relations. It is very important to trace how these relations influence the design characteristics as the design features of the object ultimately have a direct impact on the cost estimates (see Fig. 4.8). The logical modeling of the studied life cycle stage of the cost estimate's object (block 1.2) includes a solution to the following basic problems: constructing a model for the internal structure of the process and reflecting the basic developmental stages of the object in the given life cycle stage; the establishing of the basic characteristics in the state of the processes inherent to the given stage and their quantitative estimate; constructing a logical model reflecting the reciprocal influence of the state characteristics of the life cycle stage.

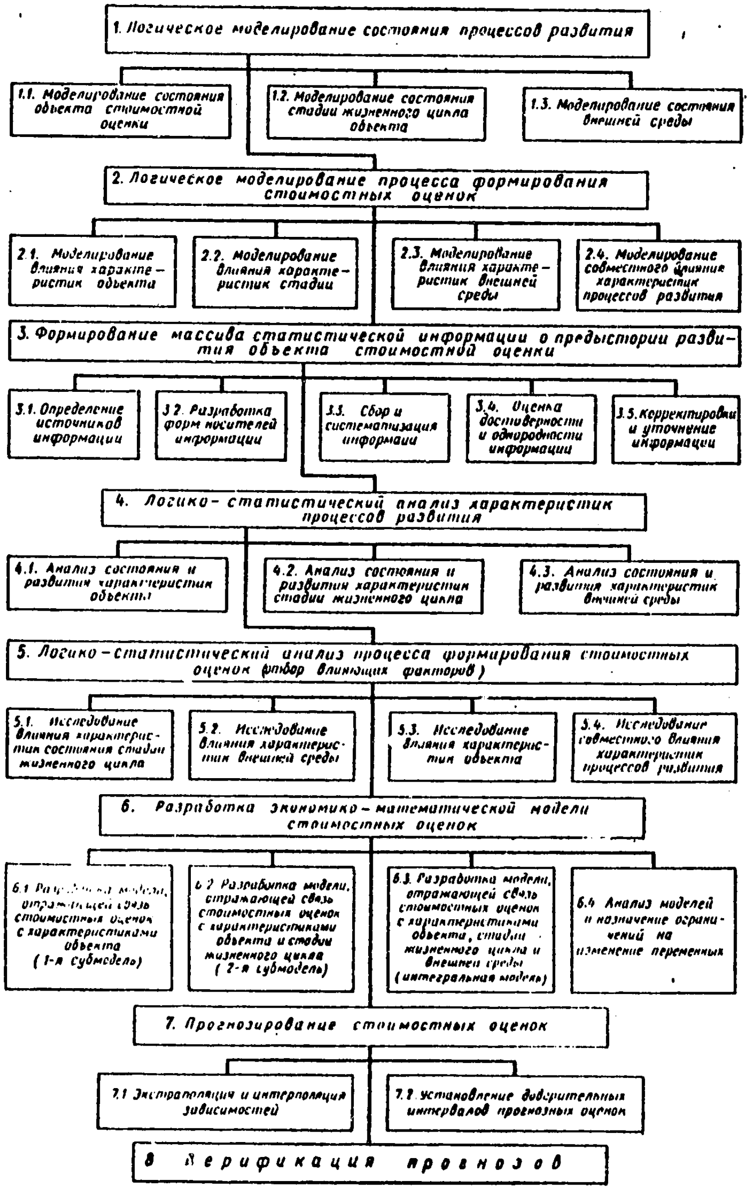


Fig. 4.9. Basic elements in a block diagram for modeling and forecasting cost estimates for the BTS functional elements

[See the key on following page.]

[Key to Fig. 4.9 on preceding page]

1--Logical modeling of state of development processes; 1.1--Modeling state of cost estimate object; 1.2--Modeling state of stage of object's life cycle; 1.3--Modeling state of environment; 2--Logical modeling of formative process of cost estimates; 2.1--Modeling influence of object's characteristics; 2.2--Modeling influence of characteristics of stage; 2.3--Modeling influence of environmental characteristics; 2.4--Modeling combined influence for characteristics of developmental processes; 3--Forming the file of statistical information on the developmental prehistory of the cost estimate's object; 3.1--Determining information sources; 3.2--Elaborating forms of information carriers; 3.3--Collection and systematization of information; 3.4--Assessing reliability and uniformity of information; 3.5--Correcting and adjusting information; 4--Logical-statistical analysis of characteristics in developmental process; 4.1--Analysis of state and development of object's characteristics; 4.2--Analysis of state and development of characteristics for life cycle stage; 4.3--Analysis of state and development of environmental characteristics; 5--Logical-statistical analysis of formative process of cost estimates (selection of influencing factors); 5.1--Research on influence of state characteristics of life cycle stage; 5.2--Research on influence of environmental characteristics; 5.3--Research on influence of object's characteristics; 5.4--Research on combined influence of characteristics of developmental processes; 6--Elaboration of mathematical economics model of cost estimate; 6.1--Elaboration of model reflecting relation of cost estimates with object's characteristics (first submodel); 6.2--Elaboration of model reflecting relation of cost estimates to characteristics of object and life cycle stage (second submodel); 6.3--Elaboration of model reflecting relation of cost estimates to characteristics of object, life cycle stage and environment (integral model); 6.4--Analysis of models and setting of constraints on change in variables; 7--Forecasting of cost estimates; 7.1--Extrapolation and interpolation of dependences; 7.2--Setting of confidence intervals for forecast estimates; 8--Verification of forecasts

The internal structure model of the process characterizing the designated life cycle stage should have sufficient detailing in order to isolate the permanent elements in the process which do not depend upon the properties and particular features of the cost estimate's object or upon the elements the composition of which varies from object to object. This will make it possible subsequently to isolate from the total expenditures that portion which is most related to the change in the characteristics of the cost estimate's object.

The logical modeling of the state of the environment (block 1.3) envisages the following: establishing the sources and nature of the external effects; establishing the characteristics of the state of the sources and a quantitative estimate of these characteristics; constructing a model which shows the links of the environment with the characteristics of the developmental processes. These problems are solved separately for the cost estimate's object and the stages of its life cycle. Such a dividing makes it possible to form a clearer notion of the mechanism of the environment's influence on the cost estimates. This ultimately facilitates the choice of the indicators which reflect this influence in constructing the cost model.

The concluding phase of the logical modeling stage is the constructing of a hypothetical model for the forming of the cost estimates in each life cycle stage of the system's functional element (block 2). For this purpose one first constructs isolated models (blocks 2.1, 2.2, 2.3) each of which reflects the influence disclosed in the previous stages on the cost estimates for the characteristics of the object, the life cycle stage and the environment.

As a result of uniting these models, an integral model is constructed for the forming of the cost estimates (block 2.4). The integral model should not be the mere total of the isolated models. In constructing it it is essential to consider the interrelations between the characteristics of the different groups of factors in forming the cost estimates.

Regardless that the integral model is to a certain degree approximate it serves as a good basis for further research. Using it it is possible to establish a preliminary list of problems which the researcher will encounter in the process of constructing the cost model. Most importantly, the logical modeling stage makes it possible to draw up a list of characteristics from which it is possible to start forming the statistical information file on the developmental prehistory of the cost estimate's object.

The formation of the file of initial statistical information (block 3) is an interation process with a direct link and feedback with all subsequent modeling stages. This stage consists of a number of repeating operations: determining the information sources, working out the forms of information carriers, the collection and systematization of information, assessing the reliability and uniformity of the information, its correcting and the clarification of the block (3.1-3.5).

The iterative nature of the process of forming the data file is explained primarily by the fact that the logical modeling makes it possible to obtain only a rough sketch of the scheme for forming the cost estimates. As a consequence of this, the data content corresponding to the initial list of characteristics is subsequently adjusted as the research is deepened. Moreover, such operations as assessing the reliability and uniformity of the information can be carried out only in the process of statistical analysis or in constructing the cost models. Due to the designated factors the formation of the initial information file can be considered complete after the final variation of the model is obtained.

In the general instance the file of initial statistical information is a time-systematized series of prototypes for the system's functional element each of which has its corresponding actual cost estimates, functional and design parameters, as well as characteristics of the studied life cycle stage and the environment. The designated characteristics reflect the state of the developmental processes for the cost estimate's object over the examined period of its prehistory. In possessing a sufficiently representative series of state characteristics it is possible to move on to studying the process of forming the cost estimates.

However, it is advisable beforehand to carry out a logical statistical analysis of the very characteristics of the state of the object's developmental processes (block 4), bearing in mind the realization of the following basic goals: establishing the integral quantitative indicators which accumulate the influence of the corresponding group of characteristics on the cost estimates; dividing the overall aggregate of characteristics into dynamic which have certain time trends and the stationary; the static which are not related to the developmental processes; the continuous and the discrete. The first is essential in order to narrow the range of examined indicators and to level out the existing contradictions between the individual characteristics; the latter is required in order to clarify the basic tasks and to choose the research methods.

The integral indicators can be obtained directly by the correlation and regression analysis methods (see Section 2.5) and on the basis of expert estimates. In the latter instance the integration of the characteristics can be attained by employing estimates for the relative weights of the individual characteristics. Here two methods can be employed for constructing the integral indicators: additive and multiplicative. In the first instance the integral indicator's model has the form of the weighted total of the individual characteristics:

$$P_\Sigma = \sum_{i=1}^{m} \beta_i p_i, \tag{4.19}$$

in the second, a weighted product

$$P_\Sigma = \prod_{i=1}^{m} p_i^{\beta_i}, \tag{4.20}$$

where p_i--the normed value of characteristics i;[2]
β_i--the relative weight of characteristics i in the process of forming the cost estimates $\left(\sum_{i=1}^{m} \beta_i = 1\right)$.

For obtaining the integral indicators it is essential that the synthesized characteristics be expressed using a finite number of quantitative measurements. However, in a number of instances the state characteristics of the developmental process reflect purely qualitative features which cannot always be reflected by a finite number of quantitative estimates. Moreover, it is essential to bear in mind that in forecasting the cost estimates on the basis of integral indicators, the need arises of compiling forecasts for each synthesized characteristic. This can cause insurmountable difficulties if special models and methods are not worked out for forecasting the individual state characteristics of the developmental processes. For this reason under certain circumstances, for simplifying the procedures of cost estimate forecasting, unformalized methods can be used for choosing the integral indicators. These methods are based upon the proposing and subsequent checking of the logical hypotheses about the relationships of the individual characteristics and their impact on the cost estimates.

In using the unformalized methods for synthesizing the characteristics, the use of the following basic rules can be helpful: the movement from cause to effect, since

---

[2] One of the possible methods for norming the characteristics is given on page 109.

the latter can be determined quantitatively; the movement from effect to cause if the same conditions are satisfied; the choice of indicators which determine certain states of the examined processes but are not related to the latter by cause-and-effect ties; the localization of the influence of factors which cannot be estimated quantitatively by the creation of isolated models; the establishing of indicators for adjusting the models. Let us illustrate the use of these rules in the given sequence using specific examples.

The characteristics of the state of the environment (block 1.3, Fig. 4.9) in which the macroeconomic factors belong, include: the level of the material and technical base of the national economic sectors; the industrial and national economic management system; the level of specialization and cooperation, the structure and principles for locating the industrial sectors.

In the process of developing the BTS, all factors progress, that is, the material and technical base of the sectors is improved, new highly productive equipment and new production processes are introduced, the level of automation and mechanization rises, the forms of the organization of production and labor, the management levels and so forth are improved. In accord with the changes in domestic and foreign policy, the state confronts the national economy with new tasks the implementing of which necessitates a reorganization of the management system. New industrial sectors arise, while the forms of specialization and cooperation, the structure of the sectors, the principles of their location and so forth change.

A quantitative estimate for the state of even one aspect of this process would require the use of more than a score indicators. Even in this instance there would be no absolute certainty that all the particular features of the influence of these factors on the cost estimates had been considered. The elaboration process and the assessing of the adequacy of the integral indicator's model require a great deal of time and effort and the model is obsolete before it has been obtained.

Moreover, its use for forecasting purposes will be of dubious value as it is essential to elucidate the state of the entire aggregate of indicators over the forecasted period and this requires the presence of forecast models for each indicator. At the same time, a change in the macrovariables can have a noticeable impact on the forecasting results of the cost estimates even over a short period of time. The way out of this situation can be found if one expresses the influence of the macroeconomic factors by the consequences of those processes the action of which they reflect.

The processes of scientific and technical development in the national economic sectors lead to a rise in social labor productivity and this ultimately is expressed in a decline in the cost of industrial products. Consequently, the introduction of an indicator for the reduction of costs into the general mathematical economics model will make it possible indirectly to consider the effect of the macroeconomic factors on the cost estimate level in the forecasting. Thus, the action of the factors which act as a cause of a certain process can be considered by using an indicator which reflects the consequences of this process.

The use of the second rule is characteristic for modeling a process involving a change in the cost estimates under the influence of the design features of BTS functional elements. Among the design features, for example, one could put: the geometric shapes of the design elements, the mechanical properties of the employed materials, the classes of precision and roughness in piece working, the design scheme of the article and so forth. Each of the designated characteristics reflects a certain aspect in the mechanism of forming the cost estimates. However, a whole series of characteristics reflects the quality features of the object and cannot be expressed by a finite number of quantitative measurements. Where this is possible, their number is extremely great. From this it follows that the attempt to synthesize such characteristics encounters as many difficulties as was the case with the environmental characteristics. At the same time, if one turns to the factors which determine the need for various design solutions, one will see that they are completely caused by the characteristics of system operational effectiveness.

Thus, an increase in the payload of a passenger aircraft involves an increase in its overall dimensions and design weight. If constraints are imposed on these characteristics, then the need arises for using lighter and stronger elements making it possible to increase the effective volume of the cargo and passenger cabins. This, in turn, involves a rise in the mechanical properties of the structural materials, it complicates the configuration of the design elements and so forth. In precisely the same way an increase in aircraft speed requires the use of heat-resistant high-alloyed steels and alloys, a change in the geometric shape, the complicating of aircraft systems and so forth.

Thus, if an aircraft's basic functional characteristics are incorporated in its cost model, one can thereby establish the influence on the cost estimates of the design features which are the consequence of a change in the functional characteristics.

The third of the formulated rules for selecting the indicators is widely employed in analyzing the microeconomic factors. With the overall development level of the material and technical base, specialization, concentration and cooperation of the serial-production and developing enterprises, the state of the production processes for the functional elements is largely determined by the scale of their serial output.

For example, with a high general level of mechanization for the production processes at a specific enterprise, the mechanization level for the manufacturing processes of the cost estimate's objects can be below the average due to the small scale of output. The larger the scale of output for the functional elements the more preferred it is to deepen specialization, increase the level of automation, the equipping of the production processes and so forth.

Thus, the scale of production, without being linked to the designated factors by cause-and-effect ties, determines the possible states of the production processes for the individual subsystems. Consequently, the incorporation of indicators for the output scale into the model will make it possible to reflect the impact of the corresponding group of factors on the expenditures level when the modeling of the effect of their direct indicators causes certain difficulties.

If the influence of the individual factors is only qualitative, it is possible to resort to isolated models. This is how they proceed when one functional element combines the functions of several special elements in the system. For example, a booster-cruise engine combines the functions of the take-off and main engine. With the same characteristics such a combined engine will differ from an ordinary cruise engine in a number of design and technological features and many of these features cannot be given a uniform quantitative estimate. In this instance, a mathematical economics model is worked out separately for each engine subclass: lift-off, cruise and booster-cruise.

The last of the named rules is employed in those instances when the influence of the factors is a discrete one, that is, certain conditions influencing the structure and level of the cost estimates change in abrupt shifts. Thus, in converting to the new planning and economic incentive system adopted by the September Plenum of the CPSU Central Committee in 1965, there were changes in prices for many types of raw products and materials, the sources and forms of paying bonuses to industrial workers and so forth. In this instance the mathematical economics models obtained on the basis of retrospective information and encompassing the period preceding the changeover of the enterprises to the new system would not be able to reflect the particular features of that period for which the cost forecasts were being made. From this viewpoint, forecasting accuracy could be increased by incorporating in the model correction factors which would consider the corresponding changes in the cost estimate structure.

From what has been stated it can be seen that the stage of selecting the quantitative indicators which reflect the influence of factors that determine the basic formative patterns of the cost estimates is linked to carrying out profound comparative and semantic analysis of information about the actual expenditures, to studying the particular features of the extant prototypes of the object and to studying the organizational-economic conditions of their creation, production and operation as well as the environment characteristics. The results of this complex of research are represented by a set of hypotheses on the nature of the impact of the designated factors on the cost estimates. The advanced hypotheses require an experimental verification.

The hypotheses are verified by the methods of mathematical statistics and probability theory. This stage is called selecting the influencing factors (block 5). Below we give the procedure for selecting the influencing factors for constructing cost estimate models in the form of multiple regression equations.

The choice of the influencing factors consists in establishing a certain range of quantitative indicators which during the prehistory period of the object's development had a determining influence on the process of cost estimate formation. Here it is assumed that the prototypes of the functional element represent a particular selection from a certain general aggregate of elements the state characteristics of which are distributed normally relative to their averages $\bar{X}_1$, $\bar{X}_2$, $\ldots$, $\bar{X}_p$, and that the spread of these properties can be described by the variances σ_1^2, σ_2^2, $\ldots$, σ_p^2, where p--the number of indicators describing the entire aggregate of studied state characteristics.

If the expenditures in the general aggregate are distributed normally with the average $\bar{X}_1$ and the variance σ_1^2, then the essential influence is considered to be the one of that indicator out of the total set X_2, X_3, $\ldots$, X_{p-1}, the variance of which explains a certain portion of the general variance σ_1^2. However due to the absence of data on the general aggregate, under real conditions one operates with the estimates of the averages $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_p$ and the variances $S_1^2, S_2^2, \ldots, S_p^2$ calculated on the basis of sampling data.

In this regard an estimate is introduced for the essentialness of the observed statistical relations, that is, an estimate of to what degree the relationships established on the basis of sampling research reflect the state of the general aggregate.

The estimates of essentialness or reliability are made with a certain predetermined degree of confidence in the correctness of the advanced hypotheses. This confidence is expressed numerically by the probability that our estimates will be within the limits of a certain confidence interval the width of which depends upon the amount of the mean square deviation S, the size of the sampling n and the set probability $\mathcal{P}$, and does not depend upon the shape of the distribution curve of the sampling data [30]. In particular, the true mean of the general aggregate is estimated by the interval

$$\overline{X} - t_{\mathcal{P}} \frac{S}{\sqrt{n}} < \overline{X} < \overline{X} + t_{\mathcal{P}} \frac{S}{\sqrt{n}}, \tag{4.21}$$

where $t_{\mathcal{P}}$--the criterion for estimating the significance of the random value $\overline{X}$ with the set level of fiduciary probability $\mathcal{P}$) and the number of degrees of freedom $\nu = n - 1$.

Formula (4.21) is read as follows: with a probability equal to $\mathcal{P}$, it is possible to assert that the mean general aggregate $\overline{X}$ is within the limits $\overline{X} \pm t_{\mathcal{P}}(S/\sqrt{n})$. The use of the significance estimate criteria plays a large role in studying the process of expenditure formation as it makes it possible to isolate the regular from the random.

Before moving on to an estimate of the quantitative effect of the indicators chosen on the basis of preliminary analysis, it is essential to make certain that the examined aggregate of objects is uniform from the viewpoint of the quality features which can be described by quantitative indicators. For example, it is essential to establish how much the cost estimates are influenced by the design and production features caused by the combining of the functions of lift-off and cruise engines into one propulsion unit. In formal terms it is essential to answer the question: are the lift-off, cruise and booster-cruise engines a part of a general overall aggregate from the viewpoint of the influence of their design and production differences on the level of the cost estimates?

Since each of the designated engine varieties is characterized by certain expenditures, the task is to disclose how essential is the difference between the expenditures on each engine group. For these purposes we can employ the method of estimating the differences between average values [30].

The essence of this method is in testing the hypothesis that the two independent particular aggregates with a volume n_1 and n_2 have been taken from the same normally distributed general aggregate having a mean value $\overline{x}$ and a variance σ^2. If this is the case, then the difference between the particular values $\overline{x}_{11}$ and $\overline{x}_{12}$ should not differ substantially from zero. This so-called zero hypothesis is tested with the Student criterion t:

$$t = \frac{\bar{x}_{11} - \bar{x}_{12}}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} < t_{\mathcal{P}}, \tag{4.22}$$

where $\bar{x}_{11}$ and $\bar{x}_{12}$--average expenditures for each of the compared engine groups;
s--square root of the full estimate for the variance of the difference $\bar{x}_{11}$ and $\bar{x}_{12}$.

The criterion $t_{\mathcal{P}}$ is found from tables with a number of degrees of freedom $\nu = n_1 + n_2 - 2$. If the condition of (4.22) is not satisfied, it follows from this that the influence of the design and production features of the compared engine varieties on the cost estimates is so great that their joint study can lead to distorted ideas about the influence of the remaining factors on the cost estimate formation process.

In addition to the described method, similar problems can be solved using rank correlation methods (see, for example, [47]).

The correspondingly grouped prototype aggregates for the cost estimate object are subsequently studied for the purpose of establishing the quantitative ties between expenditures on each stage of their life cycle and the quantitative indicators reflecting the influence of the individual groups of factors.

The relation between expenditures and any indicator reflecting the influence of one or a certain aggregate of factors is established by paired correlation coefficients for linear relations and by correlation indices if the relation is nonlinear (2.84). In the latter instance it is also possible to use a correlation coefficient but with the stipulation that the natural variables are replaced by their nonlinear functions (see Section 2.5.2).

The significance of the correlation coefficients and indexes is estimated from (2.86) or (2.87). If the conditions of (2.86) and (2.87) are satisfied, the influence of the given indicator on expenditures is considered significant. However, under real conditions one must deal not with one but with several indicators. Here it is essential to bring out which of the designated indicators has had a significant impact on the expenditures during the prehistory period of the system's development.

The basic difficulty in solving this problem is that between the indicators themselves there are definite ties called a covariation of variables [47]. Thus, there is a relation between the speed and range of an aircraft, the weight and power of a radar, the specific thrust and specific weight of an engine and so forth. In all instances it is important to determine to what degree one or another indicator influences the expenditures if the remaining ones are fixed. For this purpose a partial (pure) correlation coefficient (or index) is used:

$$r_{12.34\ldots p} = \frac{r_{12.34\ldots(p-1)} - r_{1p.34\ldots(p-1)}\, r_{2p.34\ldots(p-1)}}{\sqrt{\left(1 - r^2_{1p.34\ldots(p-1)}\right)\left(1 - r^2_{2p.34\ldots(p-1)}\right)}}, \tag{4.23}$$

where $r_{12.34\ldots p}$--partial correlation coefficient between indicator x_1 and x_2 with $x_3, x_4, \ldots, x_p$ as constant.

The partial correlation index is employed in the case of nonlinear relations and is figured using the same formjla under the condition that the variables are expressed by the corresponding functions.

The significance estimate for a partial correlation coefficient or index is carried out using the Fisher z ceiterion

$$z = \frac{1}{2} \ln \frac{1 + r_{12.34 \ldots p}}{1 - r_{12.34 \ldots p}} > z_{0,05} \qquad (4.24)$$

with degrees of freedom $\nu = n - 2 - (p - 1)$ [47].

This criterion makes it possible to select out of a multiplicity of indicators those the influence of which on expenditures are determining and to begin to construct multiple regression equations (block 6).

For modeling cost estimates it is possible to use linear and nonlinear multiple regression equations of the sorts (2.119)-(2.122). Thus, cost estimate modeling necessitates the choice of a form of relation between the variables.

The nonlinearity of a majority of dependences between the cost estimates and the factors of their formation is a general logical prerequisite for choosing the type of mathematical economics cost model. This is due, in particular, to the existence of sensitivity thresholds in the cost estimates to a change in individual indicators. For example, an increase in the production scale of functional elements leads to a decline in their production cost. However the rate of this decline is not constant, since however great the scale of output costs always maintain a certain value which differs significantly from zero. Moreover, the sensitivity of the cost estimates through an increase in the functional characteristics of the system elements rises as the characteristics approach their limit amounts.

The list of arguments in favor of nonlinear models could be extended, however this is better put off until the following sections where we will examine the particular features of forming the cost estimates by the life cycle stages. But here we would add that in employing general multiple regression equations as a model of the cost estimates the nonlinear models are preferable, since they are capable of simultaneously describing linear and nonlinear relationships. For example, if the actual relation between variables is a linear one and a logarithmically linear model (2.120) has been chosen, the parameters of the variables linked to the cost estimates in a linear manner will be close to one. Thus, if in (2.120) the parameter $b_2 \approx 1$ and the model has the form

$$x_1 = b_1 x_2 x_3^{b_3} x_4^{b_4},$$

then this shows that the relationship of the variables x_1 and x_2 with the fixed x_3, x_4 is a linear one. Of course it must be remembered that here there is a certain simplification of the relationships. Moreover, what has been said cannot be extended to models (2.121) and (2.122). For this reason the final judgment about the type of dependence must be made on the basis of the statistical criteria for estimating the model after calculating the parameters of the regression equations (see Section 2.5.2).

The examination of the methods for constructing and estimating the multiple regression equations shown in Chapter 2 makes it possible to point up certain properties of them which must be considered in modeling the cost estimates.

1. The relationship between the dependent variable and each independent variable is described by the same mathematical function.

2. The number of model variables is limited to the quantity of objects for which statistical information exists (to the volume of the initial aggregate). Thus, from (2.138) it can be seen that a ratio $p \leqslant n-1$ should be maintained between the number of variables p and the volume of the initial aggregate n. Otherwise no conclusions can be drawn on the model's reliability.

3. An increase in the number of independent variables leads to a drop in the model's sensitivity to a change in each individual variable, since the overall variation of the dependent variable is broken down into ever-smaller components. From (2.130) it can be seen that the multiple regression coefficients are linearly dependent upon the correlation coefficients. The latter, even with the complete absence of a relation between the variables, are always different from zero due to the existence of unobservable estimate mistakes. As a result, with the addition of a new variable, the amount of each of the regression coefficients will diminish, since the correlation coefficients are interdependent. Moreover, if the dependence between the variables cannot be strictly linearized, the estimates of the mean variances and, consequently, the correlation coefficients are greatly biased and this leads to the distorting of the model's parameters.

The listed properties of regression equations pose the following basic problems: the choice of the configuration of the mathematical economics cost model in which the specific features of forming the cost estimates would be reflected by unique analytical functions; the reduction in the number of simultaneously modeled factors. These problems can be solved using factor modeling, that is, by constructing composite mathematical economics models of the cost estimates.

The composite structural model should consist of several submodels, each of which can function independently under the conditions of the known constancy of the other influencing factors. In this instance each submodel is constructed with certain fixed values of the indicators which are independent variables of other submodels and is expressed by a mathematical function which best corresponds to the characteristics of the described fragment of the cost estimate formation process.

The choice of the influencing factors and the modeling of cost estimates for the purpose of obtaining a composite model must be carried out in the following sequence: a study of the influence of the state indicators of the life cycle (l.c.) stage; the eliminating of the influence of local (static) relations; studying the influence of the state indicators of the environment; eliminating the influence of discrete and continuously operating characteristics; studying the influence of the object's characteristics; synthesis of the submodels and estimating the accuracy of the composite model.

In considering that the influence of the individual groups of factors is modeled separately, the composite model can consist simultaneously of simple and multiple regression equations which reflect the statistical ties, of time trend models as well as elements inherent to the comparative models. In other words, factor modeling makes it possible to employ a wide range of forecasting methods for the cost estimates. For this reason the composite model is the most flexible forecasting tool for the cost estimates as it possesses high sensitivity to the characteristics of the BTS developmental processes.

The concluding stage in modeling the cost estimates for the BTS functional elements is the setting of constraints on the changes in the independent variables. The establishing of constraints for the change of the independent variables is required to determine the acceptable limits for applying the mathematical economics models expressed by multiple regression equations in forecasting the cost estimates.

The constraints on the changes in the independent variables can be divided into internal and external. The internal constraints are imposed on changes in the independent variables within a range of observed values in the initial statistical aggregate, and the external ones are imposed beyond this range. The problem of setting the constraints is directly linked to the procedure of forecasting the cost estimates and to the calculating of errors and setting the confidence intervals. For this reason all the listed questions will be examined simultaneously.

As was pointed out, multiple regression equations are determined by breaking down the general variance of the dependent variable into components, each of which is explained by a variance of a certain independent variable. The mean measure of variability for the dependent variable, with a change in any independent variable, is characterized by the corresponding parameter of the equation. Each parameter of a multiple regression equation expresses a quantitative relation between the dependent and independent variables under the condition that all the remaining variables remain unchanged.

The designated property for the parameters of multiple regression equations is caused by the fact that in calculating the amount of a certain parameter, the influence of other variables is eliminated. The eliminating of the influence of variables is achieved by considering their joint paired distributions as is seen from the the formula proposed by E. Yule [47] for determining the multiple regression coefficient

$$b_{12.34\ldots p} = r_{12.34\ldots p}\frac{\sigma_{1.34\ldots p}}{\sigma_{2.34\ldots p}}, \qquad (4.25)$$

where $r_{12.34\ldots p}$--partial correlation coefficient determined from (4.23);
$\sigma_{2\cdot 34\ldots p}$, $\sigma_{2.34\ldots p}$--mean square deviations of dependent and independent variables.

The partial correlation index employed in the case of nonlinear relations is calculated from the same formula under the condition that the variables are expressed by their nonlinear functions.

From (4.23) and (4.25) it follows that the multiple regression coefficients maintain their force only within the areas of the joint distribution of the independent variables in the limits of the observed range of their change in the initial statistical aggregate. In other words, the domain of existence of the function expressed by a multiple regression equation is restricted to the intersection areas of the subsets belonging to the sets of the possible combinations of independent variables.

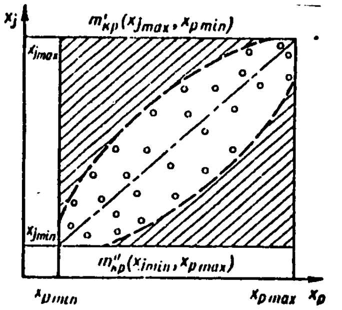


Fig. 4.10. Domain of existence for function expressed by regression equation with two independent variables with a one-way link

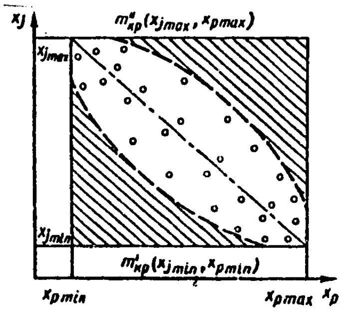


Fig. 4.11. Domain of existence for function expressed by regression equation with two independent variables with feedback

Such domains are shown in Figs. 4.10 and 4.11 for equations with two independent variables with a one-way (Fig. 4.10) and feedback (Fig. 4.11) correlation link with the following conditions:

The two variables x_j and x_p are independent variables of the multiple regression equation $x_1 = b_I + b_j x_j + b_p x_p$;

The initial statistical aggregate from the data of which the equation parameters b_1, b_j and b_p were calculated contains n pairs of values of x_j and x_{pi}, where i = 1, 2, ..., n;

The values of the variables lie within the limits set by the system of inequalities

$$x_{j\,\min} < x_j < x_{j\,\max};$$
$$x_{p\,\min} < x_p < x_{p\,\max}. \qquad (4.26)$$

The rectangles shown in Figs. 4.10 and 4.11 have been formed by the intersection of the verticles with the abscissas $x_{p\,\max}$ and $x_{p\,\min}$ and the horizontals with the ordinates $x_{j\,\max}$ and $x_{j\,\min}$. Obviously the area of the rectangles contains a multiplicity of all the possible combinations of the variables x_j and x_p within the limits set by the system of inequalities of (4.26), while the areas limited by the dotted lines contain only those combinations of independent variables which are encountered in the initial statistical aggregate. The multiple regression coefficients have been determined precisely in these areas.

From this it follows that the shaded areas of the rectanbles represent zones with undetermined (unpredictable) interpolation errors. And the greatest errors in forecasting must be expected when the point with the coordinates of m (x_j; x_p) coincide with one of the critical points m'_{kp} and m''_{kp} which are most distant from the center of of the area of the joint distribution of the variables. Thus, interpolation errors can reach very impressive amounts if the appropriate constraints are not imposed on the changes in the independent variables.

From what has been said it follows that the constraints defined by the system of inequalities of (4.26) are necessary but insufficient disciplining conditions for the interpolation forecasts.

The disciplining conditions can be considered sufficient if to the constraints of (4.26) one adds a constraint which would be an area of acceptable values for one variable when the other assumed a certain new previously unobservable value.

Such an area can be described if it is assumed that the joint distribution of two variables which are the independent variables in the regression equation x_p and x_j is a normal one.

$$f(x_p, x_j) = \frac{1}{2\pi\sqrt{1-r^2}\sigma_{xp}\sigma_{xj}} e^{-\frac{1}{2(1-r^2)}\left[\left(\frac{x_p-\bar{x}_p}{\sigma_{xp}}\right)^2 + \left(\frac{x_j-\bar{x}_j}{\sigma_{xj}}\right)^2 - 2\frac{x_p-\bar{x}_p}{\sigma_{xp}}\frac{x_j-\bar{x}_j}{\sigma_{xj}}\right]},$$

where r--correlation coefficient for values x_p and x_j.

In crossing the distribution surface with planes parallel to the plane x_p0x_j and projecting the sections on plane x_p0x_j, we obtain a family of similar and uniformly distributed ellipse with a common center $(\bar{x}_p, \bar{x}_j)$, the equations of which have the form

$$T_\alpha^2 = \frac{1}{1-r^2}\left[\left(\frac{x_p-\bar{x}_p}{\sigma_{x_p}}\right)^2 + \left(\frac{x_j-\bar{x}_j}{\sigma_{xj}}\right)^2 - 2\frac{x_p-\bar{x}_p}{\sigma_{xp}}\frac{x_j-\bar{x}_j}{\sigma_{xj}}r\right], \tag{4.27}$$

where α--the fiduciary probability.

As a general statistic which is calculated from the values of many variables, it is possible to use the statistic T_α^2, which is related to a Fisher distribution in the following manner:

$$T_\alpha^2 = \frac{2(n-1)}{n-2}F_\alpha, \tag{4.28}$$

where F_α--the statistic having 2 and (n - 2) degrees of freedom.

The radiuses of the ellipses will change depending upon the value of the fiduciary probability α. From equation (4.27) it can be seen that the ellipse is determined by five parameters $\bar{x}_p$, $\bar{x}_j$, σ_{xj}, σ_{xp}, r.

The symmetry axes of the ellipse form with the 0xp axis the angles determined by the equation

$$\operatorname{tg} 2\varphi = \frac{2r\sigma_{x_p}\sigma_{x_j}}{\sigma_{x_p}^2 - \sigma_{x_j}^2}. \tag{4.29}$$

The equation of the ellipse assumes a canonical form if the coordinate axes coincide with the symmetry axes of the ellipse. Let us designate the variables in (4.27):

$$t_{xp} = \frac{x_p - \bar{x}_p}{\sigma_{x_p}}; \quad t_{xj} = \frac{x_j - \bar{x}_j}{\sigma_{x_j}}.$$

Let us move the beginning of the coordinates to point ($\bar{x}_p$, $\bar{x}_j$) and turn the coordinate axes to angle ϕ, as determined by the equation of (4.29). In this case, the equation of the ellipse will be expressed by the formula

$$\frac{1}{1-r^2}\left(t_{x_p}^2 + t_{x_j}^2 - 2t_{x_p}t_{x_j}r\right) = T_\alpha^2. \tag{4.30}$$

In a standardized scale, the center of the ellipse is at the start of the coordinates ($t_{x_p} = 0$, $t_{x_j} = 0$) and the axes of the ellipse are directed along the bisector of the coordinate angles: the first and third angles for the first axis and the second and fourth for the second axis.

The coordinates for the ends of the first axis are:

$$A_1\left(T_\alpha\sqrt{\frac{1+r}{2}},\ T_\alpha\sqrt{\frac{1+r}{2}}\right);$$

$$A_2\left(-T_\alpha\sqrt{\frac{1+r}{2}},\ -T_\alpha\sqrt{\frac{1+r}{2}}\right).$$

The coordinates for the ends of the second axis are:

$$B_1\left(T_\alpha\sqrt{\frac{1-r}{2}},\ -T_\alpha\sqrt{\frac{1-r}{2}}\right);$$

$$B_2\left(-T_\alpha\sqrt{\frac{1-r}{2}},\ T_\alpha\sqrt{\frac{1-r}{2}}\right).$$

With $r > 0$, the first axis is the major axis of the ellipse and the second is the minor one. The greater $|r|$, the more the ellipse is extended along the major axis. If $r = 0$, that is, the random amounts x_p and x_j are not correlated, then the ellipse is turned into the circumference of the radius T_α and the equation of this circumference is $t_{xp}^2 + t_{xj}^2 = T_\alpha^2$.

The transition to the variables x_p and x_j is carried out according to the following formulas:

$$x_p = t_{x_p}\sigma_{x_p} + \bar{x}_p; \quad x_j = t_{x_j}\sigma_{x_j} + \bar{x}_j.$$

Thus, the constraints for a model with two independent variables are determined by the system

$$\left\{\begin{array}{l} x_{p\,\min} < x_p < x_{p\,\max}; \\ x_{j\,\min} < x_j < x_{j\,\max}; \\ \left(\frac{x_p-\bar{x}_p}{\sigma_{x_p}}\right)^2+\left(\frac{x_j-\bar{x}_j}{\sigma_{x_j}}\right)^2-2\frac{x_p-\bar{x}_p}{\sigma_{x_p}}\frac{x_j-\bar{x}_j}{\sigma_{x_j}}r=T_\alpha^2(1-r^2). \end{array}\right. \tag{4.31}$$

In the general case, for n variables, the equation for the function limiting the confidence area can be written in the following manner:

$$\begin{gathered} T_\alpha^2 = X\sigma^{-1}X^T; \\ X=[X_1-\overline{X}_1,\ X_2-\overline{X}_2,\ \dots,\ X_n-\overline{X}_n]. \end{gathered} \tag{4.32}$$

The matrix X^T is the transposed matrix for the X matrix

$$X^T=\begin{bmatrix} X-\overline{X}_1 \\ X_2-\overline{X}_2 \\ X_n-\overline{X}_n \end{bmatrix}.$$

The inverse matrix σ^{-1} is calculated in the following manner:

$$\begin{bmatrix} \sigma_{x1}^2\sigma_{x1x2}\cdots\sigma_{x1xn} \\ \cdots\cdots\cdots\cdots \\ \sigma_{x1xn}\cdots\sigma_{xn}^2 \end{bmatrix}.$$

Consequently, if there are n independent variables, then the constraints imposed on their change will assume the form:

$$\left\{\begin{array}{l} x_{1\,\min} < x_1 < x_{1\,\max}; \\ x_{2\,\min} < x_2 < x_{2\,\max}; \\ \dots\dots\dots\dots \\ x_{n\,\min} < x_n < x_{n\,\max}; \\ T_\alpha^2 = X\sigma^{-1}X^T. \end{array}\right. \tag{4.33}$$

In the event of a curvilinear regression, instead of a correlation coefficient a correlation ratio is used (2.84). The center of the ellipse is located at the point with the coordinates x_{pc}, x_{jc} which represent the coordinates for the center of gravity of the curve and are calculated by the formulas:

$$x_{pc}=\frac{\int_a^b x_p\sqrt{1+\left(f'_{x_p,\,x_j}\right)^2dx}}{\int_a^b\sqrt{1+\left(f'_{x_p,\,x_j}\right)^2dx}};\quad x_{jc}=\frac{\int_a^b x_j\sqrt{1+\left(f'_{x_p,\,x_j}\right)^2dx}}{\int_a^b\sqrt{1+\left(f'_{x_p,\,x_j}\right)^2dx}}. \tag{4.34}$$

The values of the integration limits correspond to the beginning and end points of the given curve. The confidence area for a curvilinear regression of two variables is shown in Fig. 4.12.

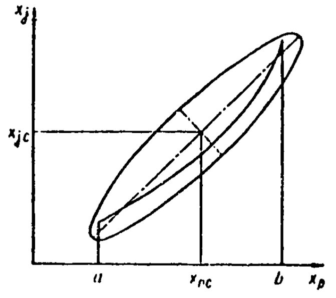


Fig. 4.12. Confidence area for curvilinear regression of two variables

With the formulated constraints on the change in the independent variables of regression models, the probability of the occurrence of unpredictable interpolation errors is minimized.

In observing the disciplining conditions expressed by the corresponding system of constraints on the changes in the independent variables, it is possible to establish the degree of accuracy of the forecast estimates using the multiple regression equation (see Section 2.5.2).

The accuracy of forecasting cost estimates is inversely proportional to the error of the individual prediction (2.140) and is determined by the confidence interval of (2.143).

As was already pointed out, the closer the new values of the independent variables come to the limits of the observed (in the initial statistical aggregate) range of changes the greater the forecast errors, since the errors of the regression coefficients are equal to zero with the equality of the independent variables to their averages. However, in the space of the observed combinations of independent variables, the nature and strength of influence of an individual variable on the cost estimates are not uniform. As this is so, the greater the danger is that the nature and strength of influence of the variable on the cost estimates will change if the variable goes beyond the limits of this space.

The question of to what degree the relations change between the cost estimate and the independent variables cannot be solved by formal methods and for this reason the imposing of external constraints represents largely a conceptual problem. Its solution depends to what degree the values of the independent variables differ from their limit states near which the probability of disrupting the established ties is increased. In this regard the imposing of external constraints becomes possible if for each variable values are set for the limit states as their existence is beyond dispute. Thus, for the BTS elements it is essential to know the theoretically achievable limits of their functional characteristics. Then, proceeding from an analysis of the time trends of the functional characteristics, it is possible to formulate the external constraints.

An indispensable condition for formalizing the external constraints should be a sufficient distance of the extreme limits of the acceptable changes in the functional characteristics from their limit states. Here an important disciplining condition should be the constraints related to the areas of the reciprocal correspondence of functional characteristics, that is, constraints of the type (4.27) should be satisfactory. In particular, in extrapolating a statistical dependence of two independent variables x_j and x_p (see Fig. 4.10), in assigning new values $x_p > x_{p\ max}$ or $x_p < x_{p\ min}$, it is essential that the extrapolation value of x_j remain within the limits of the confidence area of (4.27). In observing all the necessary precautions, the extrapolation mistakes can be commensurable with the interpolation errors on the nearby boundaries of the existence domains of functions expressed by the multiple regression equations. For calculating the extrapolation errors it is possible to use formula (2.140), and for calculating the confidence intervals, (2.143), as an assumption on the possibility of extrapolation in principle presupposes apriori that the errors of the forecast estimates are unsystematic and are subordinate to a normal distribution law.

In forecasting cost estimates using composite mathematical economics models which are a linear and nonlinear combination of statistical dependences, the total forecast errors is calculated from (2.144) and (2.145).

The forecast for the value estimates of BTS functional systems, like any other forecast, requires verification. The verification of cost forecasts is also particularly essential in extrapolation when there are fears that the established dependences can be disrupted.

The verification of forecasts for mathematical economics models can be carried out using a duplicate forecast made by a different method. For the purposes of verification of cost forecasts it is most effective to use the conversion factor method if the integral model of the mean conversion factors includes indicators which could be selected in the process of logical and statistical analysis of the formative process of the cost estimates (block 4 in Fig. 4.9). Then the mean conversion factor model and the mathematical economics model of the cost estimates will be comparable, since they will contain the same indicators. The difference in the regression coefficients from the weight coefficients set by experts will indicate the basic sources of discrepancies in the results of the forecast estimates and this can help in determining the area of search for a better model if the decision is taken to carry out a repeated cycle of analyzing and modeling the cost estimates. Thus, the use of the conversion factor method for the purposes of forecast verification, in addition to carrying out verification per se, can help clarify the mathematical economics models of the cost estimates.

In concluding an examination of the forecasting methods for cost estimates, we would like to draw attention to the following. The mathematical economics models, undoubtedly, are the most objective and flexible tools in forecasting the BTS cost estimates. In reflecting the general patterns in the change of the cost estimates under the influence of the characteristics of the system developmental processes, the mathematical economics models make it possible to assess the consequences and effectiveness of the decisions taken to control these processes. Using the mathematical economics models, along with selecting the optimum BTS parameters, it is possible to choose the variations for the processes of creating, serial production and operation of their functional elements. However, all these merits are realized only in the instance that the model has been correctly constructed in carrying out all the logical and formal procedures examined by us.

The modeling process, as one can see, entails great expenditures of time and requires the involvement of highly skilled specialists. For this reason, as was already pointed out in 4.2.1, the use of mathematical economics models should be justified by the forecasting goals. Among these goals we would put the choice of optimum parameters for systems, production processes and so forth, that is, the solving of such problems when it is important to know the influence of one or another parameter on the cost estimates. But in those instances when the influence of the individual parameters on the cost estimates is not major, the time trend extrapolation methods can be successfully employed for forecasting purposes.

As for the accuracy of the estimates, as we have seen in forecasting using the mathematical economics models, this depends largely upon how correctly the constraints have been set for changes in the variables of the mathematical economics models. Since the procedure of setting the constraints is not always formalizeable, there is always the probability of the occurrence of unpredictable errors. Thus in individual instances the extrapolation errors using mathematical economics models can be comparable with the errors of time trend extrapolation.

4.3. Basic Patterns in the Formation and the Forecasting Methods for the Costs of NIR and OKR of Large Technical Systems

Scientific research work and prototype design work (NIOKR) which are frequently linked by the common term "creation," are the two most important stages in the life cycle of a BTS. Precisely here, in these stages, start the processes of scientific and technical development of systems and these determine the evolution of the appearance of the systems and their functional properties and the means and methods of materializing these properties in the broad sense of this word. NIOKR includes a series of events in the system's life cycle from the genesis of the initial idea (concept) for creating the system up to the construction and development of the prototype.

Scientific research work (NIR) includes research on the processes of the external and internal functioning of the systems as well as the physicochemical processes occurring in the subsystems and functional elements. Along with research on the systems and their elements, the NIR is carried out for the purpose of seeking out new design materials, fuel and other energy resources, production processes and the methods of organizing, planning and controlling the creation and production of the systems.

On the basis of the tactical and technical requirements formulated considering the results of the predesign scientific research on the systems, their prototype design work (OKR) is carried out. The OKR encompasses a range of design work and the building (manufacturing) and testing of the prototypes of the systems and their elements.

The increased complexity of scientific and technical problems related to the creation of modern technology is a reason for the constant increase in the absolute expenditures on NIOKR and increasing their proportional amount in general industrial expenditures. Thus, in U.S. industry in 1940-1965, the volume of expenditures on NIOKR rose almost 200-fold while their proportional amount in the nation's budget increased from 0.82 to 15.4 percent (Fig. 4.13). Here in the total volume of allocations on NIOKR, around 90 percent is taken up by expenditures related to the creation of large technical systems such as: aviation, missile and missile-space complexes and thermonuclear weapons.

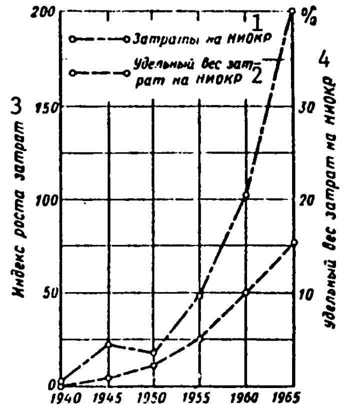


Fig. 4.13. Dynamics of expenditures on NIOKR in the United States

Key: 1--Expenditures on NIOKR; 2--Proportional amount of expenditures on NIOKR; 3--Expenditure growth index; 4--Proportional amount of expenditures on NIOKR

In the general NIOKR expenditures, the highest proportional amount is taken up by OKR. This is explained by the high material and labor intensiveness of manufacturing system prototypes, by the complexity of their testing programs and by other factors. However, a characteristic feature of recent years has been the more rapid growth rate of outlays on NIR in comparison with expenditures on OKR. The more rapid growth rate of NIR expenditures has been an objective pattern of the present-day scientific and technical revolution, since the scientific potential which ensures the increased rates and the continuity of scientific and technical development is created mainly as a result of scientific research.

The increased outlays on the NIOKR for large technical systems, in outstripping the growth rates for NIOKR expenditures in general industrial outlays, gives important significance to the questions of forecasting production costs for the NIOKR of the BTS. The methods of forecasting the production costs of scientific research differ substantially from the forecasting methods of the cost estimates employed in the remaining stages of the BTS life cycle. This is determined both by the adopted practice of calculating actual expenditures on the NIR as well as by certain characteristic features of sectorial NIR.

NIR is carried out by sectorial scientific research institutes (NII). The end product of the NII is the solution to a certain scientific problem related to a rise in the operational effectiveness of the systems, to an improvement in the functional characteristics and design of the system elements and to an improvement in the processes of creating and producing the systems and the methods of planning, organizing and managing these processes. Along with theoretical research, the NII also carry out experimental work which is done for the purpose of testing the results of theoretical research on mock-ups and models.

In terms of its character the scientific research work conducted by the NII is divided into three types:

1) Fundamental research consisting in the solving of broad general theoretical problems related to the creation of the system as a whole or a range of uniform articles comprising various aircraft systems;

2) Exploratory (preliminary) research conducted in the aim of disclosing the possibility and advisability of solving various problems at the given moment and choosing the most rational areas of research;

3) Applied research aimed at solving particular problems involving an improvement in the quality of designs, production processes, the organization of production for a certain type of product and so forth.

The exploratory, fundamental and applied research create the scientific potential for carryingout OKR and serial production of the systems and their functional elements. Here a majority of the NIR results is used in working out and producing a series of system generations. Thus at each moment of time the NII are solving problems related to the development prospects of the sector. The designated diversity and perspective orientation of the NIR are the reason that the attempts at modeling NIR expenditures, depending upon the characteristics of the systems or the individual stages of their life cycle, have not been crowned by success.

At present the methods of an indirect estimate for the costs of NIR have become widespread. An indirect estimate for the cost of the NIR related to the creation of BTS presupposes a forecasting of these expenditures proportional to a certain cost estimate which is sensitive to a change in the system characteristics. Considering that scientific research by its nature comes closest to the processes of OKR and, in addition, is financed from the same source, the state budget, the NIR expenditures are set proportionately to the OKR costs. Here the share of expenditures on sectorial NIR which should be put against the costs of the OKR of a specific system is determined from an analysis of the existing proportions in the sectorial budget allocations for the creation of new technology.

In the general instance, expenditures on the creation of a system functional element are determined by the formula

$$C_{niokr} = C_{okr}(1 + K_{nir}), \tag{4.35}$$

where C_{okr}--the costs of the OKR for a system element;
K_{nir}--a proportionality factor characterizing the ratios existing in a given sector between NIR and OKR costs.

As was already pointed out, the growth rates for the NIR and OKR expenditures are not constant over time. This is reflected in the fact that the proportionality coefficient K_{nir} has a certain time trend and for this reason the forecasting of NIR expenditures with the known C_{okr} presupposes the modeling and extrapolation of a time trend for the proportionality factor.

If a function reflecting the time trend of a proportionality factor is differentiable, for forecasting the expenditures it is possible to use the following formula [34]:

$$K_{nir} = K_{nir}^{0} \frac{dK_{nir}(\tau)}{d\tau} \Delta\tau, \tag{4.36}$$

where K_{nir}^{0}--the index for the NIR and OKR cost ratio in the base period τ_0;
$\frac{dK_{nir}(\tau)}{d\tau}$--the gradient for the index of the NIR and OKR cost ratio;
$\Delta\tau$--the time gap; $\Delta\tau = \tau - \tau_0$.

The correct determining of the possible ratios for the NIR and OKR expenditures in the future has a marked impact on the results of the forecast estimate for the cost of creating the BTS. For this reason it is essential to show great attention to an analysis of the time trends of the proportionality factor K_{nir} in order that the function describing these trends correctly reflects the general patterns in forming the NIOKR expenditure structure.

However, there is no doubt that the accuracy of forecasting the OKR costs for the BTS has the basic impact on the results of the forecast estimate, as the OKR holds the largest proportional amount in the general expenditures on creating the systems. The model for the OKR costs of a BTS functional element can generally be expressed by a multiple regression equation in which the independent variables represent the characteristics of the functional element and the process of its OKR. For a logarithmically linear form of dependence, a model for OKR costs is written

$$\tilde{C}_{okr} = b_1 \prod_{j=2}^{m} x_j^{b_j}. \tag{4.37}$$

The model (4.37) is a very approximate reflection of the process involved in forming the cost of an experimental subject. A multiple regression equation does not consider, and indeed cannot consider, all the particular features of OKR and these particular features, as will be shown below, have a substantial impact on the process of OKR cost formation.

The OKR of systems and their functional elements is carried out at prototype design bureaus (OKB) and is characterized by three basic stages: designing, the manufacturing of prototypes, testing and adjustment.

The first stage includes the work of designing the prototype, carrying out experiments and working out the working drawings and technical documents required for manufacturing and testing the prototypes. In the second stage work is done to manufacture the prototypes, as well as to design and manufacture special fittings and tools. The third, concluding stage of the OKR provides for the carrying out of experimental adjustments and testing for both the system as a whole as well as its individual elements.

Each of the designated stages is carried out to a certain degree by an independent functional complex of the OKB. Designing is carried out by the designing complex (PKK), the manufacturing of ptototypes by the production complex (PK) and testing by the testing complex (IK). For this reason the OKR process for a system element can be represented as the process of the sequential transformation by each functional OKB complex of its specific inputs into specific outputs (Table 4.7). The specific input of the PKK is the information input, for the PK it is the material, and for the IK the object (the latter in Table 4.7 is shown as part of the material input).

The specific information input is the concept of the system (element) represented in the form of a technical designing requirement or tactical-technical requirements (TTT) for the system (element). This is the organizing specific input of the OKB. In the designing process, in addition, scientific-technical information is employed

Table 4.7

Inputs & outputs	Name of inputs and outputs of BTS development				
	information	material (object)	energy	personnel	disturbances
Inputs	Development technical requirement (TZ). Scientific-technical information on characteristics of analogous type systems. Soviet & foreign patent information. Data on new materials equipment, production processes, etc. Plan and directive documents	Raw products, basic materials, semifinished products, preassembled articles. Subsystem elements to replace worn out ones. System prototypes delivered by cooperation, their individual aggregates	Various types of energy. Elements of buildings, structures. Production equipment, tools, fittings, attachments, instruments to replace worn out ones	Personnel trained to work at enterprise	Changes in TZ. Nonfulfillment of TTT by subsystem developers. Change in plan quotas. Cancelling of financing. Violation of deliveries, State Standards, etc.
Outputs	Design documents for newly created system. Descriptions, instructions, testing results statements. TTT for subsystems. Directive-set production methods, design changes & other production documents. Reports, statements, proposals on promising developments	Prototype (standard) of system. Models of individual elements of complex. Mock-ups & models. Examples of fittings and attachments	Loss of energy into environment. Loss of elements of buildings, structures, production equipment, fittings instruments, etc., as consequence of wear. Production losses.	Departure of personnel	Deviation from the TTT for system. Exceeding development dates. Overexpenditure of resources. Incomplete design. Poor quality design documents, etc.

which forms the basis of knowledge in this area. The embodiment of the concept in the design of a system (element) occurs by mental processes which generate new information in the form of design decisions. The variation of the plan which satisfies the TTT forms the information specific output of the OKB.

The specific input for the production complex is the material input in the form of the subjects of labor which change their properties under the impact of the means and implements of labor. The prototypes of the system (element) are the result of this effect and they form the specific object output of the PK.

The specific input of the testing complex is the system (element) with its initial level of ambiguity relative to the conformity of the functions and the ability for them to meet the TTT. This ambiguity is minimized or completely surmounted under the effect of the range of monitoring and metering equipment employed in the testing process. A system which carries out the specified functions with the required level of effectiveness forms the specific object output of the IK and the OKB as a system.

The distinction of the specific inputs and outputs determines the uniqueness of the basic OKR stages and this must be considered in forecasting the OKR expenditures for BTS. The predominance of information processing the generating processes gives the design stage an exploratory nature and this explains the high degree of ambiguity in the process and its end result. The ambiguity of a design process is particularly great in the early stages of elaborating the system's design. These include the elaboration of the preliminary project, the technical requirement and the technical proposal.

A technical development requirement gives the technical, operational and production requirements made on the system and its subsystems. A technical requirement establishes the basic purpose, the flight-technical characteristics of the article to be developed, the conditions for its employment as well as the composition and basic characteristics of the subsystems and elements.

A technical proposal is an aggregate of design documents which contain the technical and technical-economic feasibility studies for the elaboration of the system on a basis of analyzing the technical requirement and the different variations for the possible solutions for the articles as well as a comparative estimate of the solutions considering the design and operational features of the to-be-developed and existing articles as well as patent materials.

In the above-listed stages, the design studies are particularly closely tied to the NIR, as in the process of the preliminary studies new problems are frequently brought to light the solving of which necessitates special NIR. The results of this research cycle are considered in the further working out of the design.

The degree of ambiguity in the result is noticeably reduced only in the draft design stage. The amount of information in this stage increases significantly and the data from the preliminary design stage are clarified and established by experimental data. Experimenting is conducted by creating mock-ups of the individual subsystems and elements in the system.

A draft study is the first stage in which the design parameters of the article to be designed are defined and the design appearance of the system to be designed is formed. On the basis of the draft design, in the process of technical designing, the working components of assembly units, schematic diagrams for fuel supply, electrical equipment and so forth are worked out.

In the stage of technical designing, questions arise which are analogous to the questions of draft designing, but the number of variations for solutions is substantially reduced, since some of them were rejected as a result of mock-up construction in the draft design stage.

The stage of working designing is characterized by an extensive work front to create the drawings for the article, its individual units, assembly units and parts. On the basis of the working drawings, directive methods are elaborated for manufacturing the prototype of the system (element).

But still, regardless of the rather thorough elaboration of the design, the ambiguity about the conformity of the system (element) to the TTT remains high until the carrying out of full-scale testing, and for this reason in a number of instances the failures discovered in the testing lead to the halting of experimental subjects even before they are complete. The reason for the premature halting of OKR at times can be found in the miscalculations made in elaborating the preliminary project and the TTT for the system. The low scientific and technical potential contained in the TTT leads to the obsolescence of the sytem which is in the OKR stage. Obviously under the conditions of a high degree of development ambiguity, the mistakes leading to the halting of work are one of the common patterns in the OKR. For this reason, in forecasting the expenditures on the OKR of the BTS it is essential to consider the estimate for the average probability of successfully completing the OKR.

The average or mean probability of success μ is determined by the ratio of the cost of the successfully completed work over the past period of time $\tau_k - \tau_0$ to the total cost of all the work performed over the same period [34]:

$$\mu = \frac{\int_{\tau_0}^{\tau_\text{к}} C'_{\text{окр}}(\tau)\, d\tau}{\int_{\tau_0}^{\tau_\text{к}} C_{\text{окр}}(\tau)\, d\tau}. \tag{4.38}$$

In the event of the continuing of the OKR following the testing results, the design documents and prototype undergo the corresponding changes and the testing is repeated. The number of such cycles is rather difficult to predict, and for this reason the testing process, like the designing of the system, is characterized by a high degree of ambiguity. This significantly complicates the process of modeling OKR costs for systems and their elements.

The process of manufacturing the prototypes is more determined in comparison with the first and third stages, in a number of instances it possesses common traits with the stage of serial production but also has a number of specific features.

In contrast to the first and third OKR stages, the process of prototype manufacture on the input and output has mainly material flows. Prototype production requires raw products, materials, semifinished goods and preassembled articles but its end result is the prototypes of the articles being designed.

Experimental enterprises are classified in single-unit type of production, although they differ from classic single-unit production in a certain focus of the production process and specialization of production on a uniform group of articles. In addition, in single-unit production there is no change in the technical specifications for the order and no supplementary work and this is characteristic for experimental or prototype production. The limited range and scale of output also tell on enterprise size. Experimental plants are significantly smaller than serial-production plants in terms of the number of employees and the productive capital.

The manufacturing of new articles in units (or experimental batches) causes a number of particular features in experimental production such as the lower equipping of production with special tools and fittings, and, consequently, the small capacity of the tool shops; the use in the production process basically of universal equipment, the high skills of production workers and the consolidated elaboration of production methods.

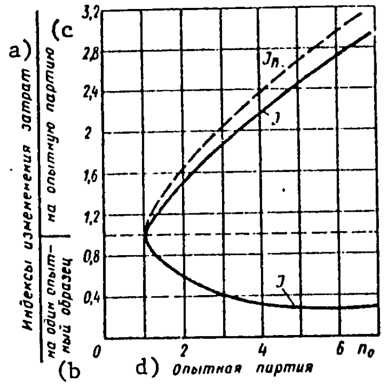


Fig. 4.14. Dynamics of expenditure indexes for developing articles depending upon size of experimental batch

Key: a--Indexes for change in expenditures; b--For one prototype; c--For an experimental batch; d--Experimental batch

The designated features of experimental production increase the production cost of prototypes in comparison with costs in serial production. However, the dynamics of expenditures on manufacturing articles in the process of experimental and serial production shows common trends: the costs of each subsequent specimen are less than the previous one. In other words, in experimental production costs are influenced by the degree of developing the design and the manufacturing methods of the article. This is one of the most important features of serial production (this question will be examined in detail in the following section). As a consequence of the influence of the degree of production development on the costs of prototypes, the costs of the experimental batch increase more slowly than the sizes of an experimental batch. The latter tells also on the behavior of OKR expenditures with a change in the number of prototypes.

The cost dynamics of prototypes J_1 of the experimental batch J_n and the total OKR expenditures for articles J, depending upon the number of examples manufactured from the start of prototype production, are shown in Fig. 4.14. The designated patterns are apparent even with the comparatively small experimental batches. In a number of instances the experimental batches reach significant sizes and then the similarity of experimental production with serial production is further increased.

The entire amount of OKR [34] is carried out on the first prototype complex which includes the flying prototype, the models for static and dynamic testing and repeated load testing (for each subsequent flying model, starting with the second, there is only the stage of manufacturing and testing with the necessary amount of rework and adjustment). In this manner consideration is given to the invariance of design expenditures to the size of the experimental batch. A model for this type of OKR costs has the following configuration:

$$\tilde{C}_{\text{окр}} = \frac{1}{\mu}(\gamma_0 + \gamma_1 n_0^{\lambda_{n+1}})\,\tilde{C}_{\text{окр}_1}(\bar{x}), \tag{4.39}$$

where γ_0--the proportional amount of conditionally fixed expenditures in the costs of the first experimental complex;

γ_0--the proportional amount of variable expenditures in the costs of the first prototype set;

n_0--the number of examples in the experimental batch;

λ_{n+1}--elasticity coefficient for variable expenditures in relation to size of exterimental batch determined empirically;

$\tilde{C}_{okr\,1}(\bar{x})$--the costs of the first experimental set of aircraft expressed in the form of an equation of dependence upon the vector of the functional and design characteristics.

The specific weights of the conditionally fixed and variable expenditures are determined on the basis of analyzing the time trends in the OKR cost structure over the previous period of time.

The model (4.39) is most effective for forecasting the OKR costs of the system elements when the number of examples in the experimental batch is comparatively small and their purpose is controlled by the adopted testing system as occurs in the development of aircraft. In manufacturing prototypes in large batches and with significant fluctuations in the batch size (the latter often depends upon the degree of originality and newness of the articles), it is essential to consider the impact of the scale of experimental production on prototype costs. For this purpose it is essential to make the process of forming expenditures in the stage of manufacturing the prototypes into an independent object of modeling.

In considering that the number of prototypes directly or indirectly influences the testing costs, the OKR cost model is best shown in the form of the total of the particular expenditure models for each stage.

The cost model for the experimental batch generally is expressed by the following formula:

$$\tilde{C}_{\text{оп}} = C_{n_0}(\bar{x},\ n_0)\,n_0, \tag{4.40}$$

where C_{n_0}--the average cost of the prototype expressed in the form of a dependence upon the characteristics of the functional element and the number of prototypes.

The dependence of prototype costs upon the size of the experimental batch n_0 can be approximated by the step function

$$C_{n_0} = a_1 n_0^{\lambda_n}, \tag{4.41}$$

where α_1 and λ_n--the equation parameters determined empirically and $\lambda_n < 0$.

This dependence can be expressed in a dimensionless form if as the base one selects the amount of the prototype costs found from (4.41) with a certain fixed size of the experimental batch. In the given instance the most suitable base is the prototype cost corresponding to the arithmetic average from the amount of the experimental prototype batches of the functional element:

$$\bar{n}_0 = \sum_{i=1}^{n} n_{0i}/n, \tag{4.42}$$

where n--the number of prototypes in the initial statistical aggregate.

Then the relative influence of the experimental batch on prototype costs is expressed by the index

$$J_{C_{n_0}} = C_{n_0}/C_{\bar{n}_0}, \tag{4.43}$$

where $C_{\bar{n}_0}$--the cost of a prototype from a batch equal to $\bar{n}_0$.

In substituting the values C_{n_0} and $C_{\bar{n}_0}$ calculated from (4.41) for n_0 and $\bar{n}_0$, respectively, in (4.43) we obtain the dependence of prototype costs upon the size of the experimental batch in a dimensionless form

$$J_{C_{n_0}} = \left(\frac{n_0}{\bar{n}_0}\right)^{\lambda_n}. \tag{4.44}$$

Now, having expressed the influence of n_0 by $J_{C_{n_0}}$, (4.40) can be written in the following form

$$\tilde{C}_{on} = C_{n_0}(\bar{X})\, n_0 \left(\frac{n_0}{\bar{n}_0}\right)^{\lambda_n}. \tag{4.45}$$

Thus, the cost model for the OKR of functional elements which are characterized by large ranges in the change of the scale of experimental production, can be represented by the following equation:

$$\tilde{C}_{\text{окр}} = \frac{1}{\mu}\left[C_d(\bar{X}) + C_{n0}(\bar{X})\, n_0 \left(\frac{n_0}{\bar{n}_0}\right)^{\lambda_n} + C_t(\bar{X},\ n_0)\right], \tag{4.46}$$

where C_d--designing costs;
C_t--testing costs.

The establishing of the dependence of design costs C_d upon the characteristics of the functional elements, as was already pointed out, is a difficult task due to the high degree of process ambiguity at this OKR stage.

For the system functional elements where among the characteristics it is possible to isolate a base parameter that characterizes the geometric dimensions of the element (for example, the weight of the aircraft frame, absolute engine thrust and so forth), designing costs are approximately expressed by the equation of dependence upon this parameter. However the choice of the base parameter is possible not for all the functional elements of the systems while the modeling of designing costs using other parameters requires a further breaking down of the designing process into smaller items, since the sensitivity of designing costs to the various parameters is not the same for all the elements in the internal structure of this process. Thus, the obtaining of a good model for designing costs for a rather complicated functional element can take up a good deal of time and effort even for an experienced research collective.

At the same time, an analysis of the OKR expenditure structure shows that, regardless of the tendency for an increase in designing costs for complicated systems, their proportional amount in overall OKR costs is relatively slight and for certain functional elements is 2-5 percent. Naturally, under these conditions, even major errors in design expenditure forecasting will not have a substantial impact on the accuracy of the overall OKR cost estimate. For this reason, in a number of instances, when the obtaining of a reliable mathematical economics model of forecasting costs requires the carrying out of a complex range of research, it is possible to permit a certain simplification. For example, as an adequate model for the forming of designing expenditures one can adopt the average designing costs calculated from actual expenditures for the designing of the prototypes of functional elements:

$$\bar{C}_d - t_{\mathcal{P}} \frac{S_{\bar{C}_d}}{\sqrt{n}} \leqslant \tilde{C}_d \leqslant \bar{C}_d + t_{\mathcal{P}} = \frac{S_{\bar{C}_d}}{\sqrt{n}}, \qquad (4.47)$$

where $S_{\bar{C}_d}$--the estimate of the mean square deviation of actual expenditures from the arithmetic average.

The procedure for forecasting OKR costs can be further simplified if the proportional amount of designing expenditures $\tilde{\gamma}_{Cd}$ in the total OKR expenditures of a functional element is sufficiently stable or a certain time trend for this indicator is known. Then the OKR costs are determined from the formula

$$\tilde{C}_{\text{окр}} = (\tilde{C}_{\text{оп}} + \tilde{C}_t) \frac{1}{1 - \tilde{\gamma}_{C_d}}. \qquad (4.48)$$

The forecasting of testing costs C_t is somewhat facilitated by the fact that in the testing process a large amount of energy resources is required. For this reason, with other conditions being equal, the testing expenditures will be sensitive to a change in the capacity of the energy sources and to the consumption of energy resources per unit of capacity. However, due to the ambiguity of the testing process, testing cost models usually introduce a large error factor into the total error of an OKR cost forecast in comparison with the expenditure models for experimental production.

The OKR processes for the BTS are constantly being improved and this is one of the manifestations of the factor of an increase in social labor productivity. Moreover, this factor influences the costs of the products of past labor consumed in the development of the systems. In this regard, in forecasting the OKR expenditures, it is essential to consider the trend in the decline in OKR costs related to the growth of social labor productivity. The influence of this factor can be considered by using the cost time trend coefficient in the sectors producing the BTS and their functional elements:

$$K_{\tau} = (1+0.01\overline{W})^{T_e-T_o} \qquad (4.49)$$

where $\overline{W}$--the average annual reduction in the costs of industrial product consumed in the OKR process, %;

T_e--the year of carrying out the OKR of the system's elements which are the objects of the cost estimate;

T_o--the year of drawing up the cost forecast.

The total cost of working out the system S is the total expenditures on working out the individual system elements considering the possible use of results from the development of elements in other systems and the overall development cost of the system:

$$\tilde{C}_{\text{окр}\,S} = \sum_{i} \tilde{C}_{\text{окр}\,j}/\gamma_{\text{окр}\,j} + \tilde{C}_{\text{оп}\,S}, \qquad (4.50)$$

where $\tilde{C}_{okr\ j}$--the OKR costs for element j of the system;

$\gamma_{okr\ j}$--the proportional amount of expenditures on developing element j related to the development costs of the system being designed;

$C_{op\ S}$--the costs for the general development of the system.

4.4. Basic Formative Patterns and Methods for Forecasting the Costs of Serial Production of the BTS and Their Functional Elements

4.4.1. Particular Features of the Serial Production Process and Cost Formation

The process of serial production for the functional elements of the BTS can be conditionally divided into three basic stages: 1) the development of the first serial models or the stage of production preparation; 2) the development of serial production or the development stage; 3) full-scale serial output or the stage of established serial production.

The first stage encompasses the period from the start of the preparations for producing the article up to the output of the first serial model. At this stage organizational and technical measures are carried out related to preparing the enterprises to turn out the new article. Production preparations for the new article include the following measures: the rearranging and reorganization of existing shops and sections; the elaboration of serial-production drawings for the product and production fittings; the elaboration of serial production methods; the manufacturing of the required amount of production fittings (the first stage of fittings); the testing of the initial materials; the manufacture and testing of individual structural elements and units of the article; the manufacture, assembly and testing of the first serially produced model.

It is essential to point out that a portion of the work related to production preparations for the new articles is carried out in parallel with their serial output. The latter is explained by the need to shorten the overall production preparation cycle in the aim of accelerating the output of the first articles. The output of articles of the head series and their subsequent putting into operation are needed to eliminate design shortcomings the discovery of which is possibly only under ordinary operating conditions. As a consequence of this the final adjustment of the article's design is made at a serial-production enterprise in the process of disclosing and eliminating the existing shortcomings. The designated circumstances necessitate the manufacturing of the prototypes under conditions where the production process for manufacturing the article is equipped with a minimum range of special production fittings and without which the output of the article is essentially impossible.

Thus, the conditions for turning out the first serial-produced models of the article have the following particular features: production instability and a relatively low technological level; low equipping of the production processes; imperfect forms for organizing production and the work areas; the absence of work skills for the workers and insignificant experience of technical personnel; the absence of technical standards for labor intensiveness; the incompleteness of the article's design and so forth.

The listed characteristic traits inherent to the moment of completing the first serial production stage are the reason for the relatively high costs of the first articles. Subsequently, as serial production is developed, article costs decline substantially and in a number of instances by the end of the second stage are 15-20 percent of the initial.

The decline in costs at the production development stage is achieved by introducing measures aimed at raising the organizational and technical level of production. At this stage the serial production drawings and the serial production methods are finally elaborated; the manufacturing of the production set of fittings is fully completed (the level of equipping in a number of instances reaches 90-95 percent); the production areas and lines are determined; technical standards for labor intensiveness are introduced (the proportional amount of technical standards reaches 70-75 percent); adjustments and improvements are incorporated in the article's design; the workers gain work skills and the technical personnel gains experience in manufacturing the new article.

The production development stage is characterized by an increase in product output per unit of time. By the end of this period, the enterprises reach a steady production program and for this reason, in speaking about developing the production of BTS functional elements, it is essential to bear in mind not only the process of developing the design and the production methods for the articles but also reaching designed output scale.

Thus, the production development of new articles is a dual process. On the one hand, the production development of new articles is accompanied by a rise in the organizational and technical level of serial production (the qualitative aspect), and on the other, in keeping with the development of serial production there is an increase in the quantity of articles produced per unit of time (the quantitative aspect). The first factor influences primarily the labor and material intensiveness of the article. The action of the second factor is manifested in the fact that with an increase in product output per unit of time there is a decline in the share of the shop and general-plant expenditures and in the expenditures on carrying out special testing and setting-up outlays. Expenditures decline on production fittings per unit of article.

The two mentioned aspects of the process of cost formation are interrelated and the basic is the quantitative aspect. Thus, the production scale has a direct impact on the optimum level of outfitting the production processes in manufacturing the articles. Under the conditions of producing large batches of articles, the outlays on production fittings become more advisable, as here the increase rate in the expenditures on serial production provided by the increase in production outfitting outstrip the growth rate of the expenditures on outfitting. The higher the scale of product production the more the optimum level of outfitting is felt and thereby better prerequisites are created for producing the expenditures on the serial production of the articles.

The output scale ultimately predetermines the possibilities of organizing mechanized and automated production, the use of specialized production equipment and the introduction of advanced production processes.

The product output scale also has a substantial impact on the level of specialization and cooperation. The deepening of specialization and the widening of the interdepartmental and intrasectorial ties become effective only under the conditions of large-series production making it impossible to employ the most progressive means and methods of specialized production.

The scale of output, in characterizing, in addition, enterprise production capacity, is an indirect indicator of the production concentration level in the sectors specialized in producing the BTS functional elements.

The dynamics of cost reduction and the reaching of the designed product output scale are shown in Fig. 4.15. The degree of the serial production development of the design and the reaching of the designed product output scale influence not only the level and dynamics of the cost decline but also the ratio of the individual expenditures included in its structure. Fig. 4.16 shows a typical change in the cost structure occurring under the influence of the designated factors. As is seen from the graphs, in the process of the production development of the article there is a rise in the proportional amount of expenditures on wages and materials and a decline in the share of the shop and general plant outlays and the direct aggregate expenditures (expenditures on special fittings, testing and so forth) in the full production costs.

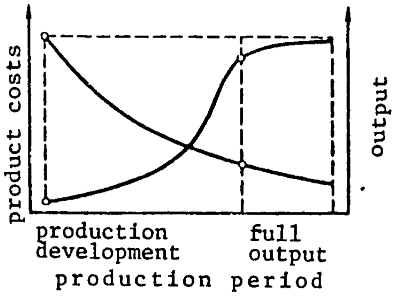


Fig. 4.15. Change in product costs and output in the process of serial production

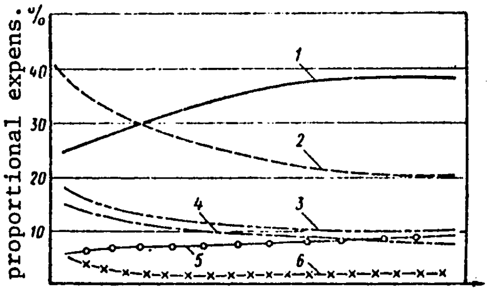


Fig. 4.16. Change in proportional amount of basic expenditures in full product costs in the process of developing their serial production:

1--Materials; 2--Shop expenditures; 3--Special fittings; 4--General plant expenditures; 5--Wages; 6--Testing

The third stage or the stage of established serial production is characterized by the following basic features: a further improvement in the design and production methods of the articles; the broadening and improving of the production set of equipment and fittings; the improving of the organization of production and labor; a rise in the automation and mechanization level of the production processes; the extensive introduction of rational methods for producing the initial stock and so forth. The given range of organizational and technical measures is carried out over the entire serial production period. During this period one most strongly feels the impact of the factors related to the rise in social labor productivity.

In the stage of full-scale serial production, when the reserves for cost reduction brought about by the newness of the article and by other particular features of the production development stage have been basically exhausted, a further decline in production costs occurs as a consequence of improving the technical means and organizational forms of serial production and the other factors which determine the growth of social labor productivity. However, in contrast to the second stage of serial production, during this stage the intensity of the cost decline does not exceed 4-2 percent per quarter.

Research on the particular features of serial production and the nature of the change in expenditures during its various stages has made it possible to draw the following conclusion. The process of forming the costs of the BTS functional elements is shaped under the influence of three basic factors which characterize the organizational and economic conditions of serial production of the systems: the organizational and technical level of production; the degree of the production development of the design and the serial output of the articles; the production scale of the articles.

The organizational and technical level of production accumulates the influence of both the macro- and microeconomic factors and the impact of the latter on the cost of new articles to a significant degree is determined by their output scale. In this regard, the task of establishing the indicators which characterize the scale of product serial output moves to the forefront.

The scale of serial output can be estimated using the following basic indicators: the number of articles manufactured from the start of serial production N; average daily output calculated for a certain production period $\bar{q}$; product output per unit of time (quarter or day) reached by a certain moment of time in the designated period q.

The number of articles N manufactured since the start of production and the average daily output $\bar{q}$ characterize the scale of serial production only when combined with the third indicator, the length of the period t during which N articles have been produced or an average daily output $\bar{q}$ has been reached. The daily output q, in contrast to the first two indicators, is an independent characteristic of the output scale at the given concrete moment of time during any segment of the serial production period of the articles.

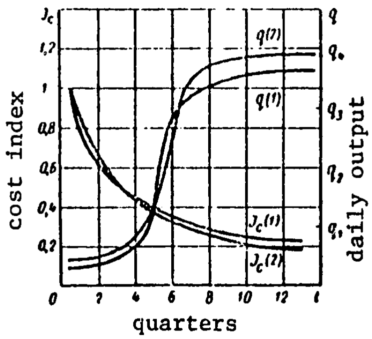


Fig. 4.17. Influence of absolute product output scales on cost dynamics and level

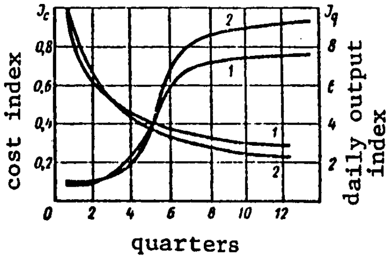


Fig. 4.18. The influence of the ratio of the initial and established product output scales (1) and (2) on their cost dynamics

The influence of daily output on the cost level and dynamics is illustrated by the examples of Figs. 4.17 and 4.18. Fig. 4.17 characterizes the dependence of cost dynamics and level upon the absolute serial output scale under the conditions of established production. In Fig. 4.18, cost dynamics are compared with the product growth rate. From the diagrams it is obvious that product costs decline more intensely the higher the scale of established production and its increase rate in the process of reaching serial output. The more intense rates of cost decline with an increase in the difference between the initial and established output are explained by the fact that the output of the first serially produced examples, independently of their quantity, occurs under approximately equal organizational and technical production conditions. Conversely, under the conditions of established serial production the organizational and technical level will be higher the greater the output scale under the same conditions.

Thus, the process of cost formation can be conditionally divided into two parts: the formation of cost dynamics and the formation of the cost level.

Cost dynamics are chiefly determined by the particular features of developing serial output while the level is determined by the output scale of developed articles and by their design and production features which are determined by the operational effectiveness factors. By the cost level one understands the cost of articles corresponding to the start of established serial production. It is felt that under these conditions the influence of the factors characterizing the processes of reaching serial production has been virtually exhausted and that subsequently the process of serial production will occur with an established level of production equipping and other organizational and technical characteristics of serial production. The start of established serial production is usually linked with the moment of producing the so-called serially developed product. Here it is assumed that the cost of a serially developed product, in contrast to the cost of the previously produced articles, shows a relative resistance to a change in the output scale and the other characteristics of the product production development processes.

Establishing the cost of a serially produced product is essential for achieving comparability in terms of the degree of serial production development for prototype articles of a system's functional element when these comprise the initial statistical aggregate. The costs of articles which are compared from the viewpoint of the degree of developing a design and the reaching of the designed output scale are used as the initial base in studying the influence of the functional and other characteristics of a functional element. In addition, the costs of a serially produced article make it possible to establish the influence of its level and output scale under the conditions of established production. Determining the costs of a serially produced article comes down to setting the moment of moving from the production development stage to the stage of full-scale serial production. The time interval (ordinarily the ordinal number of a quarter, starting from the beginning of serial output) at which this transition is made has been given the name of the "serial output point."

Above it was stated that the process of reaching serial production of products has a quantitative and qualitative aspect. Here the quantitative aspect which is the reaching of the designed output scale largely determines the qualitative aspect of the production development process, that is, a rise in the organizational and technical level of production. The conclusion arises that it is possible to speak about the moment of transition from the production development stage to full serial production only proceeding from the dynamics of daily output. Obviously the time interval by which daily output reaches a relative stable level will correspond to the serial output point.

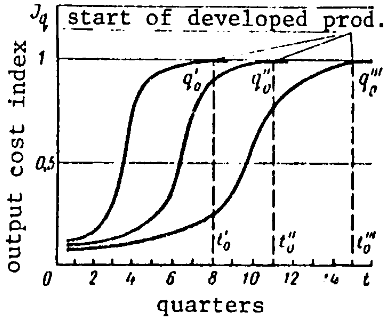


Fig. 4.19. Influence of development rate of output production program on duration of serial development period

The relative stabilization of daily output is achieved at various moments of time depending upon the pace of developing the production program. The development pace of the production program for turning out various articles can change in a rather broad range. Fig. 4.19 shows the most characteristic curves for the change in the daily output of functional elements of a system. From Fig. 4.19 it can be seen that the time intervals corresponding to the moment of relative stabilization in daily output vary between 8 and 15 quarters counting from the start of production. This shows that the determining of the serial output point is a necessary condition for eliminating the influence of the degree of product development on costs.

The serial output point, depending upon the pace of developing the production program, can be determined if one imposes certain constraints δ on the relative rate of increase in daily output per unit of time.

The time trend for daily output is best described by the dependence

$$q(t) = d_1 \operatorname{arctg}(d_2 t + d_3) + d_4, \tag{4.51}$$

where t--the time counted from the start of serial output (in quarters);
d_1, d_2, d_3, d_4--equation constants determined empirically.

The relative increase rate in daily output can be determined as

$$q'(t) : q_{max}, \tag{4.52}$$

where $q'(t)$--the derivative function of (4.51); $q'(t) = d_1 \dfrac{d_2}{1+(d_2 t+d_3)^2}$;
q_{max}--the function maximum of (4.51).

The function maximum is found from the condition

$$\lim_{t \to \infty} q(t) = d_1 \operatorname{arctg} \infty + d_4, \tag{4.53}$$

hence

$$q_{max} = d_1 \frac{\pi}{2} + d_4. \tag{4.54}$$

In reducing (4.52) to the set value of δ and solving the obtained equation ($q'(t):q_{max} = \delta$) for t, we determine the ordinal number of the quarter corresponding to the beginning of developed production, according to the following formula:

$$t_0 = \frac{1}{d_2}\left[-d_3 + \sqrt{-1 + \frac{2d_1 d_2}{\delta(\pi d_1 + d_4)}}\right]. \tag{4.55}$$

The concrete amount of the constraint δ imposed on the relative increase rate of daily output from the viewpoint of the comparability of various articles from prototypes which are in the initial statistical aggregate is not of fundamental significance. It is important that it is constant for all the compared articles. However, this amount should be rather small in order that the serial output point is to the right of the area of intense output growth. In the given example (see Fig. 4.19), the serial output point was set at $\delta = 0.01$. It is not difficult to see that under these conditions the growth of daily output is virtually halted.

The setting of a serial output point makes it possible to select with sufficient soundness the values for the costs C_0 and the daily output q_0 for serially produced articles and this plays a major role in forecasting the costs of BTS functional elements. Cost forecasting usually starts by determining the probable amount of costs for a serially produced article. After this the possible cost changes are determined over the entire extent of the serial production of the articles, that is, a forecast is made for the cost dynamics. However, research on the cost formation processes must start by analyzing the cost dynamics of articles in the process of their development and for this reason it is most convenient to start an exposition of the cost forecasting methods from the cost dynamics forecasting methods.

4.4.2. Forecasting Production Cost Dynamics of BTS Functional Elements

The forecasting of cost dynamics consists in determining the possible changes in the costs of a specific article depending upon the assumed conditions of developing its serial production. Above it was pointed out that a decline in costs is a characteristic feature in the stage of developing the production of products. However, this reduction occurs at different rates. The establishing of the regular changes in the cost reduction rate also creates the necessary prerequisites for forecasting cost dynamics.

From the preceding material (see 4.4.1), it is obvious that developing a new functional element represents a process of the adaptability of an enterprise to the serial output of the product. The qualitative aspect of this process is the developing of the design per se and its quantitative aspect is reaching the designed output scale. These are interrelated aspects. And the primary one is the quantitative aspect since the possibilities (and often the advisability) of increasing the organizational and technical level of production are restricted to the designed output scale. Proceeding from this, the methods of forecasting cost dynamics are based on the assumption that there is a quantitative relationship between the cost reduction rate and the product output scale. For this reason the methods given below differ only in the principles for estimating this relationship.

Fig. 4.20 shows the graphs for the increase in the total number of articles manufactured since the outset of serial production and the corresponding graphs for the change in costs. As is seen in the figure, product costs decline more intensely than total product output grows. In the general instance this dependence can be expressed by the multiple regression equation:

$$C = b_1 N^{b_2} t^{b_3} \prod_{j=4}^{p} x_j^{b_j}, \qquad (4.56)$$

where N--the total number of articles produced since the start of series production;
t--the time over which N articles were produced;
x_j--characteristics of the system's functional element;
$b_1, b_2, \ldots, b_p$--equation constants determined empirically.

The model of (4.56) expresses the dependence of product costs upon factors determining both cost dynamics and level. A characteristic shortcoming of the model (4.56) is that it does not meet the demand of product comparability in terms of the degree of serial production development. As was shown (see 4.2.3), the parameters of the multiple regression equation are calculated under the conditions of eliminating the influence of other variables and for this reason express the dependence between the dependent and independent variables when all the remaining variables are held on the level of their averages. This means that the constants $b_4, b_5, \ldots, b_p$ express the cost dependence of a system's functional element upon the element's characteristics with N and t corresponding to their average arithmetic values $\bar{N}$ and $\bar{t}$.

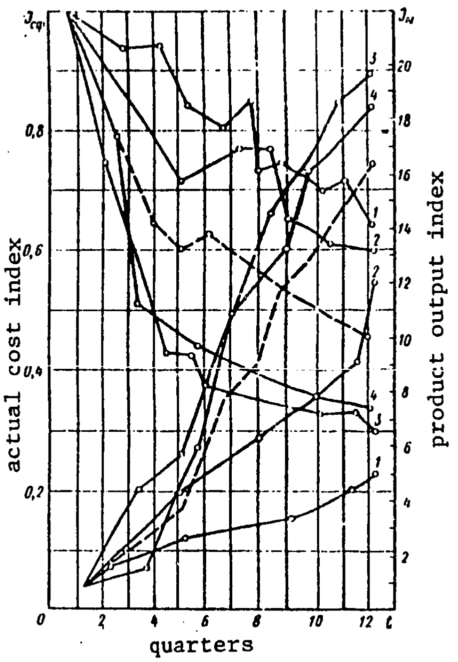


Fig. 4.20. Dynamics of product costs depending upon output size

However, as is seen in Fig. 4.19, the conditions of comparability for each individually taken article change very substantially (in the designated example t_0 changes by almost double) and depend not upon the absolute amount of N but rather upon the rate of developing output. Thus, the conformity of $\overline{N}$ and $\overline{t}$ to the comparability conditions for even one article of the initial statistical aggregate can be the result of only a random coincidence and in any event cannot be extended to all remaining articles. The given circumstance can lead to major distortions of the product cost dependence upon product characteristics, as these dependences will be influenced by the degree of serial production of the article. Naturally it is very difficult to judge the reliability of forecasts made using models of the type (4.56).

A more attractive method from the designated viewpoint is one based upon the use of a multiple regression equation to model the cost index JC_0 which represents the ratio of the article's cost observed in each time interval of the studied period of serial production to the cost of the serially produced article. In this instance the requirement is observed of product compatibility in terms of the degree of serial production and a model describing the cost dynamics has the form

$$JC_0 = b_1 N^{b_2} t^{b_3}. \qquad (4.57)$$

The model (4.57), in comparison with (4.56), provides more dependable forecasts, however the effectiveness of its use is reduced as a consequence of the existence of internal constraints on the change in the independent variables which are inherent to multiple regression equations (see Section 4.2.3).

In terms of the problem of forecasting cost dynamics using the model (4.57), to the ordinary constraints examined in Section 4.2.3, one must add the specific constraints related to the transformations of the initial statistical aggregate. These occur in calculating the parameters of the multiple regression equation of (4.57).

The initial statistical aggregate consisting of n prototype articles of a functional element contains n pairs of values (C_1, N_1); (C_2, N_2), ..., C_k; $N_k)$ for each time interval of serial production (t = 1, 2, ..., k). At the same time for calculating the parameters of a multiple regression equation it is essential that for each element C_t of set S_C there be the corresponding fully determined value of N_t for set S_N in each time interval t. For achieving this correspondence it is essential to average the values of C_1, C_2, ..., C_k and their corresponding values of N_1, N_2, ..., N_k in terms of the number of articles n in the initial aggregate for each t (t = 1, 2, ..., k).

Graphs for the time change in the averaged values C̃ and Ñ in Fig. 4.20 are depicted with a broken line. It is not hard to see that the aggregate consisting of nk values of C and N is transformed into a series where for every value of Ñ there is a uniform corresponding averaged value of C̃. If one considers that the pairs of series $(N_t, C_t)_i$ (i = 1, 2, ..., n) differ in terms of the length of the period for starting up serial production t_0, the averaging of N and C must be carried out within a certain predetermined value of this period, for example, the average duration of the period for starting up serial production for the studied aggregate of articles:

$$\bar{t}_0 = \frac{1}{n}\sum_{i}^{n} t_{0i}.$$

Under these conditions, in considering that the function C̃ = F(Ñ, t) is determined in an area bounded by an ellipse which is described by the equation

$$T^2_\alpha \quad \frac{1}{1-\eta^2}\left[\frac{(\tilde{N}-\bar{\tilde{N}})^2}{\sigma^2_{\tilde{N}}} + \frac{(t-t_C)^2}{\sigma^2_t} - 2\frac{\tilde{N}-\bar{\tilde{N}}}{\sigma_{\tilde{N}}}\frac{t-\bar{t}}{\sigma_t}\eta\right],$$

all combinations of N and t not falling into this area are in the zone of undetermined interpolation errors.

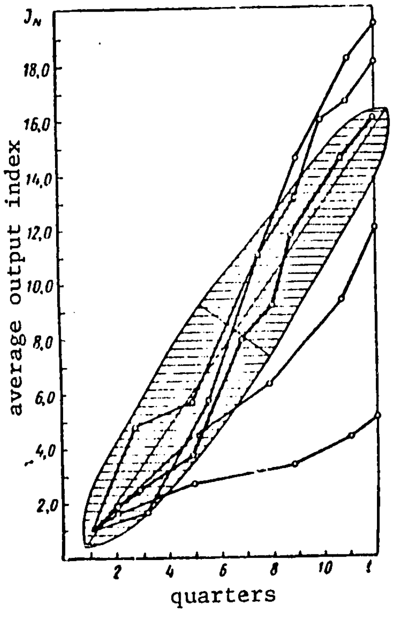


Fig. 4.21. The domain of existence of the function C̃ = F(Ñ, t)

In Fig. 4.21, the domain of existence of the function C̃ = F(Ñ, n) has been shaded. It is not hard to see that the zones of the undetermined interpolation errors comprise a significant portion of the area of the rectangle which includes the possible variations for the distribution of output over time.

From what has been said it follows that from the viewpoint of forecasting, the pair of series $(\tilde{N}_t, \tilde{C}_t)_i$ in its nature in no way differs from any other pair $(N_t, C_t)_i$, i = 1, 2, ..., n, and for this reason the equation (4.57) does not reflect the general pattern of the change in costs for the entire initial statistical aggregate.

Two basic conclusions follow from an analysis of the models (4.56) and (4.57):
1) the modeling of the dynamics and level of costs must be carried out separately;
2) the model of cost dynamics should be constructed under the conditions of the constancy of one of the two indicators N, t characterizing the production scale of the articles.

Considering that output is a controllable indicator for the production scale, it is most advisable in forecasting to utilize the influence of the latter on cost dynamics within a certain period of serial production which is constant for the studied articles. For the purposes of meeting the requirement of product comparability in terms of the degree of their serial production, the studied aggregate can include only those articles for which the duration of the production development stage t_0 varies within insignificant limits (one or two time intervals). If the cost of each article reached by the end of the selected period T and the corresponding output N_T is taken as one, then the dependence of the cost index $J_{C_{N_T}}$ upon the amount of output for an individual article is expressed by the equation

$$J_{C_{N_T}} = a_N N^{\lambda_N}; \ (\lambda < 0). \tag{4.58}$$

In a logarithmic system of coordinates, the dependence of (4.58) for a certain aggregate of articles will be represented by a family of parallel straight lines (if we disregard the random errors in estimating the λ parameter), as is shown in Fig. 4.22. From Fig. 4.22 one can see that each straight line has an intersecting point with the coordinate axes. The point of intersection for the straight line with the ordinate axis corresponds to the maximum amount of the cost index with N = 1 and is numerically equal to the parameter a_N in the equation (4.58). In turn, the point of intersection of the straight line with the abscissa axis corresponds to the cost index under the condition that the number of articles N manufactured since the start of production numerically equals the amount of the output batch N_T over period T. It becomes obvious that for each amount of the batch N_T there is a corresponding ordinate $(\lg J_{C_N})_i = a_{N_i}$ and this also characterizes the cost reduction rate.

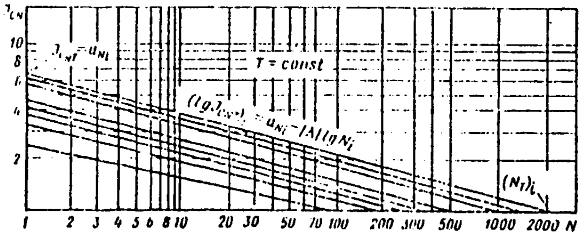


Fig. 4.22. Dependence of cost dynamics upon quantity of articles manufactured since the start of serial production

The dependence of the ordinate a_{N_i} upon $(N_T)_i$ is expressed by the equation

$$a_N = g_1 + g_2 \lg N_T. \tag{4.59}$$

In substituting (4.59) into (4.58), we obtain a model of cost dynamics in a final form:

$$J_{C_{N_T}} = (g_1 + g_2 \lg N_T) N^{\lambda_N}, \tag{4.60}$$

where λ_N = const. with T = const.

Thus, if we know the size of the product batch N_T manufactured over the period T since the start of serial production and the cost of N_T article C_{N_T}, then from (4.60) it is possible to determine the cost of any N article from this batch.

Calculations have shown a good convergence of the cost dynamics estimates for certain functional elements of systems (the deviations of the actual values from the calculations have not exceeded 10 percent). However, it is essential to emphasize that the designated method of forecasting cost dynamics entails certain complications in obtaining models of the type (4.60). Thus, in considering the constraints related to the requirement of product comparability in terms of the degree of production development, the initial statistical aggregate of articles is broken up into groups with close values of the serial output point.

Consequently, instead of one general model for cost dynamics it is necessary to build several particular models for the individual product groups. With the general smallness of the initial aggregate, the number of products in each group can be insufficient for establishing reliable statistical dependences (see Section 4.2.3). In problems involving the modeling of the costs of system functional elements where a characteristic feature is a limited number of prototypes, the latter circumstance can play a crucial role since the very possibility of constructing a model for product cost dynamics is in jeopardy.

The designated shortcomings are largely eliminated if one uses a method of cost dynamics forecasting based upon the following principle: a reduction in costs will be more noticeable the higher the output of articles per unit of time established toward the end of the stage of developing serial production. The forecasting of cost dynamics using this method is carried out from a model which is often called a "development curve" equation:

$$C(t) = \beta'_1 t^{\beta'_2}, \tag{4.61}$$

where t--the time since the start of production measured in quarters;
β'_1, β'_2--equation constants set empirically and $\beta'_2 < 0$.

From (4.61) it can be seen that the approximate value of the cost C(t) in the given quarter t can be determined by the simple substitution of a numerical value for t, if the constants β'_1 and β'_2 are known.

An analysis of (4.61) shows that with t = 1, the value of the β'_1 parameter will numerically equal the amount of product costs in the first quarter C_1 and which coincides with the start of serial output. At the same time, if in (4.61) we substitute the forecast value of the cost for the serially produced article C_0 corresponding to the serial output point t_0, then it is possible to write:

$$\tilde{C}_0 \quad \beta'_1 t_0^{\beta'_2}. \tag{4.62}$$

In making a simple transformation of (4.62), it is possible to express the β'_1 parameter through the cost of a serially produced article C_0 and the value of the serial output point:

$$\beta'_1 = \tilde{C}_0 t_0^{-\beta'_2}. \tag{4.63}$$

In substituting (4.63) in (4.61), it is possible to construct a model for forecasting product costs in any time interval of the serial production period:

$$\tilde{C}(t) = \tilde{C}_0 t_0^{\beta_2'} t^{\beta_2'}. \tag{4.64}$$

From (4.64) it can be seen that when $\tilde{C}_0$ and t_0 are correspondingly determined, for calculating the entire cost dynamics series it is merely a question of calculating the absolute amount of the exponent β_2' since for each product the values C_0 and t_0 are constant.

As was shown in Subsection 4.4.1, the cost reduction rate is influenced by the scale of established serial production (see Fig. 4.17). Research has shown that there is a close connection between the output scale expressed by the daily output of a developed product q_0 and the cost reduction rate and these are described by the absolute amount of the exponent in the equation (4.61). This relation is depicted by the following dependence:

$$|\beta_2'| = \omega_1 + \omega_2 \log q_0, \tag{4.65a}$$

where $|\beta_2'|$--the absolute amount of the exponent in the equation (4.61).

Another variety of the model (4.65a) is the dependence of β_2' upon the logarithm of the ratio of daily output for the developed product q_0 to the daily output in the first quarter of serial production q_1:

$$|\beta_2'| = \omega_1 + \omega_2 \log \frac{q_0}{q_1}. \tag{4.65b}$$

Regardless that the ratio q_0/q_1 is a simplified description for the growth rate of the daily output within the interval $(t_0 - t_1)$, the model (4.65b) does not provide any substantial improvements in the accuracy of the estimate in comparison with (4.65a). This is explained by the fact that regardless of a certain fluctuation in the initial daily outputs q_1, the first serially produced models are manufactured under approximately equal organizational and economic conditions and the degree of further change in these conditions depends upon the output scale of established serial production. For this reason, cost dynamics determined by changes in the organizational and economic conditions in the process of developing serial production with a sufficient degree of approximation is characterized by the output of the serially produced products and this is reflected by the dependence (4.65a). Thus, out of the two models (4.65a) and (4.65b) one must choose the first as the simplest one which also requires a smaller amount of information about the characteristics of the serial production process for the forecast object.

Thus, considering (4.65a), the serial production cost model which reflects the general patterns in the formation of product cost dynamics in the process of developing serial production can be written in the following form:

$$\tilde{C}(t) = \tilde{C}_0 t_0^{\omega_1 + \omega_2 \lg q_0} t^{-(\omega_1 + \omega_2 \lg q_0)}. \tag{4.66}$$

The use of the model (4.66) for forecasting involves, as one can note, the necessity of preliminary calculation of the length of the production development stage t_0 and the amount of daily output for the developed product q_0. From equations (4.55) and (4.51) it follows that for determining t_0 and q_0 it is essential to possess data on the distribution of daily output over the time of serial production q(t). However, with the existing practices of long-range planning, the enterprises are given quotas for the output of the entire batch of the given articles with a distribution of the production program over the years of the assumed period of serial production. Consequently, the quarterly distribution of daily output, as the proposed cost forecasting method requires, must be specially calculated using the data of the annual product output plans.

This problem is solved in the following manner. At the end of each year of serial production, the amount of the integral (in a running total) daily output is established using the formula:

$$Q_T = \frac{1}{D_K} \sum_{1}^{T} N_T, \tag{4.67}$$

where D_k--the average number of workdays in the quarter;

N_T--the planned output of articles for the corresponding year of serial production.

Here, the integral daily output achieved by the end of each year of serial production is assigned a corresponding number of the quarter counting from the start of serial output. For example, the integral output Q_{II} at the end of the second year is identified with Q_8 for the eighth ordinal number of the quarter. Considering that the stage of developing serial production will last usually not more than 4 years, it is enough to calculate the integral outputs for 4, 8, 12 and 16 quarters. The calculated values of the integral daily outputs are approximated by the function

$$Q(t) = \int_{t}^{T} q(t)\,dt. \tag{4.68}$$

In solving (4.68) using (4.51), we obtain the following expression for the integral function of daily output:

$$Q(t) = d_1 d_2 \left\{ (d_2 t + d_3) \text{ arctg } (d_2 t + d_3) - \frac{1}{2} \ln \left[1 + (d_2 t + d_3)^2 \right] \right\} + d_4 t. \tag{4.69}$$

If it is assumed $d_2 = 1$, then d_3 will equal the abscissa of the bending point of the daily output curve. Then the parameters d_1 and d_4 can be found in the following manner. Having designated in (4.69) the expression in braces by F(t), we have $q(t) = d_1 F(t) + d_4 t$. In using the least square method we find:

$$d_1 = \frac{\sum_{1}^{t} Q_t t - d_4 \sum_{1}^{t} t^2}{\sum_{1}^{t} F(t)\, t}; \tag{4.70}$$

$$d_4=\frac{\sum\limits_{1}^{t} Q_t F(t) \sum\limits_{1}^{t} F(t)\, t-\sum\limits_{1}^{t} F^2(t) \sum\limits_{1}^{t} Q_t t}{\left[\sum\limits_{1}^{t} F(t)\, t\right]^2-\sum\limits_{1}^{t} F^2(t) \sum\limits_{1}^{t} t^2}, \tag{4.71}$$

where t--assumes successive the values 4, 8, 12, and 16 while Q_t respectively equals $Q_4 = Q_1$; $Q_8 = Q_{II}$, ..., $Q_{16} = Q_{IV}$.

Then the calculations are carried out by the successive approximation method by substituting in (4.70) and (4.71) the values of d_3 expressed by numbers of a natural series. The criterion for the closeness of the function (4.69) to the set values of Q_t can be the maximum of the correlation index

$$\eta_Q=\sqrt{1-\frac{\sum\limits_{1}^{t}[Q_t-Q(t)]^2}{\sum\limits_{1}^{t}[Q_t-\bar{Q}]^2}}. \tag{4.72}$$

The calculations continue until a value is found for d_3 whereby the correlation index η_Q assumes a maximum value.

Thus, the forecasting of product cost dynamics, after a forecast has been made for the cost of the serially produced article C_0, should be carried out in the following sequence:

1) On the basis of the long-range annual plans for product output, an integral daily output is determined for the end of each year for the first 3 or 4 years of the proposed period of serial production;

2) Using formulas (4.70) and (4.71) by the method of successive approximations to the maximum expression of (4.72), the parameters are determined for the time trend curve of daily output (4.51);

3) For formula (4.55) an ordinal number is set for the quarter t_0 corresponding to the start of developed production;

4) Using formula (4.51) by the substituting of $t = t_0$ the amount is established for the daily output of the serially produced article q_0;

5) The found values t_0 and q_0 are substituted in (4.66); the model (4.66) is transformed to the type (4.61);

6) Product costs in any quarter of the serial production period are determined by substituting in (4.61) the numerical value of t corresponding to the ordinal number of this quarter.

The calculations made by the designated method show that the closeness of the calculated values of $\tilde{C}(t)$ to the initial data is comparable with the approximation accuracy for the cost of individual products using an equation of the type (4.61). This strengthens the confidence in this method.

In conclusion it must be pointed out that the nature of the dependences for the change in the cost reduction rate upon daily output q_0 and the size of the product batch N_T [see (4.59)] fully coincides. This shows the adequacy of the models (4.59) and (4.65a) to the depicted processes.

Attention must also be drawn to an interesting feature of these dependences: the cost reduction rates grow initially more intensely and then much more slowly than the output scale. Here, as in experimental production, one feels the existence of the cost sensitivity thresholds to a change in the production scale. This circumstance formalized by the equations (4.59) and (4.65a) can be considered not only in forecasting product costs but also in the problems of optimizing a production program for product output.

4.4.3. Forecasting the Production Cost Level of BTS Functional Elements

By the serial production cost level, as was pointed out in Subsection 4.4.1, one understands the cost of a serially produced article. Since the cost of a serially produced article is considered comparable in terms of the degree of developing serial production, research on the change in these costs in the initial statistical aggregate makes it possible to judge how costs are influenced by factors which characterize the scientific and technical development processes of large technical systems.

As a whole the costs of articles[3] are most strongly influenced by their design and production features and these, in turn, are determined by the demands made on the functional and technical-operational characteristics of the articles. At the same time, the cost level is influenced substantially by the macro- and microeconomic factors. In the forecasting of product costs, of the permanent macroeconomic factors it is important to consider the factor of the rise in social labor productivity (the methods for modeling the influence of this factor have been examined in Subsection 4.3). The influence of microeconomic factors on the product cost level is chiefly explained by the difference in the organizational and technical level and the scale of serial production under the conditions of stable product output.

As was shown above, cost dynamics are caused by changes in the organizational-technical level and the scale of production within the period of the serial output of one individual article. In opposition to this, the cost level is determined by the difference of organizational and economic conditions in the serially developed production of a parametric series of functional element prototypes. Fluctuations in the cost level under the impact of microeconomic factors are determined by a change in the qualitative and quantitative characteristics of serial production. In the process of forming the level, like cost dynamics, the main role is played by quantitative characteristics, that is, by the product output scale.

The output scale has a direct and indirect influence on the cost of developed articles, determining in the last instance the organizational and technical level of serial production (this question was examined in detail in Section 4.4.1). The

---

[3] Here and below for the sake of brevity the term "serially produced" will sometimes be omitted.

direct consequence of this dual influence of the output scale on the cost of serially produced articles is the reduction in costs with a rise in daily output under the conditions of stable serial production. This reduction occurs according to the law of a step function

$$C_{P_0} = a_q q_0^{\lambda_q}, \qquad (4.73)$$

where q_0--daily output of serially produced article;
  a_q, λ_q--equation parameters set empirically and $\lambda_q < 0$;
    C_{P_0}--the cost of a unit of base parameter P_0 (proportional cost of article).

In the general product cost model the influence of the scale of developed production can be considered using the index of proportional costs

$$J_{C_{P_0}} = \left(\frac{q_{0i}}{\bar{q}_0}\right)^{\lambda_q}, \qquad (4.74)$$

where q_{0i}--daily output of article i;
    $\bar{q}_0$--daily output taken as base.

If it is assumed $\bar{q}_0 = 1$, then (4.74) will be simplified:

$$J_{C_{P_0}} = q_{0i}^{\lambda_q}. \qquad (4.75)$$

If simultaneously one considers the influence of the factor of the rise in social labor productivity (see Section 4.3), the product cost model can be represented in the following form:

$$\tilde{C}_0 = C_0(\bar{x})\, q_0^{\lambda_q}(1 + 0{,}01\overline{W})^{\tau_n - \tau_0}, \qquad (4.76)$$

where $\tilde{C}_0(\bar{x})$--the cost of the developed article, as a function of the characteristics of the forecast object;
    τ_0--the year of output of the serially produced article used as the standard representative of the initial statistical aggregate;
    τ_n--the output year of a serially produced article which is the object of the forecast.

It is not hard to see that with $q_0 - 1$ and $\tau_n = \tau_0$, the last two multipliers of (4.76) are turned to one and the cost under these conditions will be expressed by a dependence upon the characteristics of the article $C_0 = C_0(\bar{x})$. The articles which are part of the initial statistical aggregate differ not only in terms of their characteristics but also in terms of the output scale q_0. Moreover, their output occurs in different years which encompass the forecast object's prehistory. For this reason, before beginning to analyze the influence of the object's characteristics on costs, one must adjust the actual data about the cost of the initial parametric series of articles which are the prototypes for the system's functional element. The actual data are corrected using the formulas

$$C'_{f_i} = C_{f_i} \frac{1}{q_{0i}^{\lambda_q}} \quad (\lambda_q < 0); \qquad (4.77)$$

$$C''_{f_i} = C'_{f_i}(1 + 0{,}01\overline{W})^{\tau_0 - \tau_i} \quad (\overline{W} < 0), \tag{4.78}$$

where C_{f_i}--actual initial cost of article i;
C'_{f_i}--actual cost of article i corrected for actual daily output q_{oi};
C''_{f_i}--actual cost of article i additionally corrected for actual year of serial output τ_i.

From (4.77) one can see that if the actual daily output of article i is less than one, its corrected cost will be less than the initial (the output scale is more and the cost less). Conversely, if $q_{oi} > 1$, then $C'_{f_i} > C_{f_i}$ (the reduction in the output scale entailed an increase in costs). In the same manner, if the article was manufactured before the date used as the base $_o$, then under the conditions $_o$ the cost of such an article should be lower as a consequence of the increased social labor productivity $\tau_o - \tau_i < 1$, $C''_{f_i} < C'_{f_i}$ and so forth.

Practice shows that such calculations when made sufficiently correctly significantly free the researcher from distorted ideas about the degree to which costs are influenced by the characteristics of the BTS functional elements.

Research on the cost dependence on the product characteristics ordinarily start by selecting the so-called base parameter which is an indicator determining the trend in the change of costs under the influence of the characteristics of the system functional elements. The base parameter should accumulate the influence on costs of the geometric dimensions of the object, its volume, the areas of worked surfaces, the number of design elements (pieces, assemblies and units) and so forth. The base parameter should be both a design and functional characteristic of the article.

Due to its universality the indicator of the weight of the design of system functional elements has become most widespread as the design base parameter.

An increase in the geometric dimensions and weight for the design of a functional element, with other conditions being equal, leads to an increase in the labor intensiveness of manufacturing the design elements, to a rise in the consumption of materials and to an increase in the overall dimensions and weight of production equipment. It is essential to use heavier equipment, have larger areas available and so forth. All of this, naturally, leads to increased costs for producing the products.

The influence of product weight on the absolute amount of costs has been confirmed by numerous research. As a whole there is a positive correlation between costs and weight. However, the costs of products grow more slowly than the weight of a design increases. This is explained by the fact that with an increase in the weight of a product, the number of design elements per unit of weight is reduced and the average weight of one piece grows. With an increase in the weight of a piece, the area of worked surfaces also rises but the influence of this factor on product labor intensiveness is compensated for by the reduction in the total number of piece-related operations which are performed in the various stages of the production process, by the reduction in the number of assembly units of the design and the range of production equipment, by the reduction in the consumption of materials and so forth. As a consequence of this, the cost of a unit of design weight, with an increase in its absolute weight, drops. This reduction is subordinate to the following pattern:

$$C_G = a_G G^{\lambda_G}, \tag{4.79}$$

where C_G--proportional cost of article calculated per unit of design weight;
G--design weight;
a_G, λ_G--equation parameters set empirically and $\lambda_G < 0$.

The dependence (4.79) is a reflection of the influence on costs by the factor which can be conditionally called the "design scale factor." It is important to point out that such a pattern is found in the behavior of product costs with a broad range of weight changes, as a consequence of which it is not always possible to establish the dependence (4.79) from the data of the initial statistical aggregate. For this reason, if such a dependence is lacking in forecasting product costs where the weight differs substantially to either side from the observed extreme values, the influence of this indicator on proportional costs must be additionally considered, in resorting to expert estimates for help.

The weight of a design as a basic parameter must be used in estimating the cost of the functional elements when among the characteristics of the elements it is possible to isolate a more encompassing indicator which simultaneously determines both the cost and the consumer value of the articles. In actuality weight characterizes not the functional properties of a system's element but rather the payment for these properties.

Thus, the necessity of increasing aircraft payload and range necessitates an increase in the overall weight of the design. In precisely the same way, for an increase in jet engine thrust or the power of other types of engines, for an increase in computer storage capacity or the power of energy sources and so forth, one must pay by an increase in the design weight of these articles. For this reason, in practice most often one adopts as the base parameter the characteristics which simultaneously reflect the quantitative indicators (volume, area, number of design elements and so forth) and the functional properties of the article. Because of the greater information capacity of such indicators their link with product costs is closer in comparison with the design weight.

In the process of the scientific and technical development of systems, this relationship has been further strengthened, since an increase in the base parameters of a functional element related to its specific purpose is accompanied by a desire to simultaneously reduce the product's design weight. The latter is achieved by hardening the operating processes in the functioning of the elements (increasing temperature, pressure, speed and so forth) and this, in turn, necessitates the use of expensive and hard-to-work structural materials with increased mechanical properties. Along with this, demands are increased upon the precision of parts and joints, the roughness of machined surfaces and so forth. All of this taken together leads to a rise in proportional expenditures per unit of design weight.

The tendency for a change in the proportional expenditures per unit of design weight can be traced from the example of the dynamics of the proportional cost of the frame for a number of the U.S. fighters (Fig. 4.23). As is seen from the graph, it is assumed that the tendency for an increased cost of 1 kg of airframe design weight will be maintained in the future.

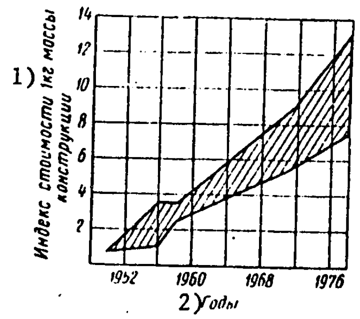


Fig. 4.23. Change in the cost of 1 kg of airframe design weight for sequential series of U.S. Air Force fighters

Key: 1--Cost index of 1 kg of design weight; 2--Years

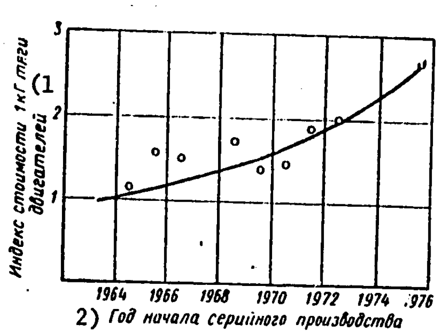


Fig. 4.24. Change in the cost of 1 kg of turbojet engine thrust over years

Key: 1--Cost index of 1 kg of engine thrust; 2--Year of beginning serial production

The increased cost of a unit of design weight for the BTS elements, in addition to increasing the indicators of their weight advancedness, is determined by the action of other, stronger operational factors of the system's specific effectiveness (see Section 4.1). For this reason, the growth trend characterizes not only the changes in the cost of a unit of design weight but also the costs per unit of the functional base parameters. This, in particular, is confirmed by research on the time trends for the cost of a unit of absolute thrust in the turbojet engines (Fig. 4.24). From the drawing it can be seen that engine cost related to 1 kg of maximum thrust shows a steady growth trend.

What has been said makes it possible to conclude that the dependence of product cost upon the functional base parameters of the products, in comparison with the dependence upon design weight, possesses greater forecasting value. However, the functional base parameter does not exhaust the influence of the entire aggregate of operational effectiveness factors on costs. For this reason, the establishing of the influence of product weight or the functional base parameters on product costs must be viewed as a preliminary stage in studying the influence of the operational effectiveness factors on costs. Subsequently, there must be a careful analysis of the behavior of proportional costs in order to establish its change patterns under the influence of the operational effectiveness factors which were not involved in the research in the first stage.

If the necessary information for these purposes is lacking or also for forecasts not requiring high estimate accuracy, the model for the time trend of the proportional costs of a functional element can be a simplified model of product costs. However, a product cost forecast which is to be used for solving optimization problems should be carried out using mathematical economics models which include product characteristics determining the growth trend of proportional expenditures as independent variables. In a general form, a proportional cost model can be represented by the following expression:

$$\tilde{C}_{P_0} = C_P(\bar{x})\, J_{C_{P_0}}, \tag{4.80}$$

where Cp($\bar{x}$)--the cost of a unit of base parameter expressing the dependence upon the object's characteristics;

J_{CP_0}--cost index for a unit of base parameter considering the influence of the "design scale factor" on costs,

$$J_{C_{P_0}} = \left(\frac{P_0}{\bar{P}_0}\right)^{\lambda_P}, \tag{4.81}$$

here P_0--base parameter of forecast object;

$\bar{P}_0$--base parameter of typical representative of initial statistical aggregate of products;

λp--equation parameter for dependence of proportional costs upon absolute amount of base parameter and found from statistical data ($\lambda p < 0$).

Considering (4.80), a general model for the costs of serially produced BTS functional elements can be represented in the following manner:

$$\tilde{C}_0 = P_0 C_{P_0}(\bar{x}) \left(\frac{P_{0i}}{\bar{P}_0}\right)^{\lambda_{P_0}} q_0^{\lambda_q} (1 + 0{,}01\overline{W})^{\tau_n - \tau_0}. \tag{4.82}$$

A model for the cost of a functional element in each time interval for the assumed period of serial production considering (4.82) is written:

$$\tilde{C}(t) = P_0 C_{P_0}(\bar{x}) \left(\frac{P_{0i}}{\bar{P}_0}\right)^{\lambda_{P_0}} q_0^{\lambda_q} (1 + 0{,}01\overline{W})^{\tau_n - \tau_0} \times t_0^{\omega_1 + \omega_2 \lg q_0} t^{-(\omega_1 + \omega_2 \lg q_0)}. \tag{4.83}$$

The procedure for constructing the mathematical economics model (4.83) is a typical example of factor modeling of cost estimates. From an analysis of (4.83) it can be seen that each element in the model reflects the influence of a certain group of factors and the uniform relations have been localized. This makes it possible to forecast costs in running through a multiplicity of variations for the possible states of the development processes and this, in turn, facilitates the procedure for choosing the optimum areas for developing the BTS.

Let us take up the possibilities of the use of the model (4.83) for forecasting the prices for new BTS functional elements. The setting of prices is an essential element in economic forecasting as the expenditures on serial production are included in the model of the system economic estimate criterion not in terms of costs but rather in the form of the prices of the system functional elements. Price forecasting involves the determining of proportional capital investments into the cost of the fixed and working capital calculated per article.

The proportional capital investments K_{ydc} into serial production undergo changes both in the process of the scientific and technical development of the system functional elements as well as in the process of developing their serial production. The direct modeling of cost estimates for fixed and working capital for price forecasting purposes requires the carrying out of special research. At the same time the trends for the change in capital investments per unit of product are largely determined by the nature of the change in the cost of the system functional elements.

As is known, the cost of fixed capital is considered in production costs by expenditures on amortization while the cost of working capital used in the creation of production inventories, incomplete production and so forth is related to a majority of the other economic elements which are components of product costs. Thus, it can conditionally be assumed that capital investments into serial production of articles will change proportionally to their cost. Consequently, in forecasting the cost of system functional elements it is enough to know the amount of the general ratio γ_f between the value of the productive capital and the cost of enterprise gross output (if at the moment of compiling the forecast the enterprise is still unknown, then for the sector) where the production of these articles will be located. Then, in knowing this ratio, the amount of proportional capital investments calculated per article can be determined from the simplified formula:

$$K_{ydc} = \gamma_f \tilde{C}(t). \tag{4.84}$$

Naturally, the effectiveness of this method for determining proportional capital investments will depend upon the stabilization of the γ_f indicator, and for this reason it is advisable to carry out research, and when necessary, the modeling of the time trend for this indicator.

4.5. Basic Formative Patterns and Forecasting Methods for BTS Operating Expenditures

4.5.1. Particular Features in the BTS Operating Process and the Classification of Expenditures for This

The operation of large technical systems includes the system functioning processes (specific operation) and the work related to the maintaining of the systems in proper technical order and functional readiness (technical operation).

Specific operation is characterized by the system functioning conditions (the interaction of the systems with the environment) and by the functional properties determined by the purpose of the systems.

Technical operation is among the auxiliary processes. Its purpose is the periodic performance of a range of measures related to the maintenance and repair of the system elements in the aim of maintaining or restoring a certain reliability and durability level in order to ensure safe operation over the service life of the functional element. Technical operation consists of maintenance and major overhaul.

Maintenance includes preventive measures performed for the purpose of maintaining and partially restoring the required level of quality, reliability and durability in the systems and subsystems. The maintenance process has a cyclical nature. The range of work involved in maintenance includes: operational maintenance (OTO) and periodic maintenance (PTO).

The basic purpose of operational maintenance is the tending and supervision of materiel and this is carried out by performing external inspections and testing the state of the systems and subsystems as well as by the partial carrying out of preventive work not requiring great expenditures of time to perform.

For maintaining the required reliability and durability in the functioning of the system over the entire period between repairs, at stipulated intervals a range of work is carried out termed periodic maintenance. The PTO work is performed as required after a certain number of system operating hours. The content, scope and time of performing the PTO are set for each functional element individually and remain constant over the amortization life. As periodic maintenance is performed, routine repairs are carried out simultaneously. Existing malfunctions are eliminated by replacing or repairing the defective parts and units.

A major overhaul is carried out for the purpose of maintaining the functional elements and the system as a whole in a state of constant readiness for use and for maintaining operating capacity between repairs during the entire service life. A major overhaul is performed after the previously set time between repairs has lapsed or with damage (breaking down) of the basic elements.

In many instances, and this is particularly characteristic for the present development stage of the BTS, a major overhaul is often combined with modernization. As a result of modernization, a functional element is returned to operation with improved performance, and for this reason in analyzing the major overhaul process it is essential to distinguish its two aspects. In the first place, a major overhaul per se as a process of restoration (eliminating physical wear) and, secondly, the performing of improvements as a modernization process, that is, the partial elimination of obsolescence in a functional element.

The operating expenditures are classified by four features (Table 4.8).

By composition all expenditure items can be divided into simple and composite. The simple expenditures are ones of a uniform nature. These include the expenditures on GSM [fuels and lubricants], amortization and current repairs. The GSM item includes expenditures only for the central element of the system, while amortization includes expenditures on renovation and major overhaul over the service life. Fuel expenditures and amortization deductions for the remaining system elements are accounted for separately. The item "Current Repair" includes expenditures on spare parts, instruments, equipment and materials consumed in current repair and maintenance as well as all expenditures on this type of work when it is performed by the enterprises.

Among the composite items are expenditures which either themselves consist of several expenditure items or include expenditures for different employee groups servicing the BTS subsystems (wages and social security deductions).

Depending upon the volume of work performed by the system, operating expenses are divided into those which depend on this volume (variable expenditures) and those which do not depend on the volume (conditionally fixed).

Table 4.8

Expenditure Items	Expenditure Classification Features							
	By composition		By method of relating to operation result		By tie with operation		By dependence upon amount of work	
	a	b	c	d	e	f	g	h
Wages		+	+	+	+	+	+	+
Deductions on social security		+	+	+	+	+	+	+
Fuels and lubricants	+		+		+			
Amortization of equipment	+		+		+		+	
Current repair	+		+		+		+	
Expenditures on ground facilities		+	+	+	+	+	+	

Key: a--Simple; b--Composite; c--Direct; d--Indirect; e--Basic; f--Overhead; g--Variable; h--Conditionally fixed

The variable expenditures include those which rise proportionally to the amount of work. For example, for aircraft systems used in the national economic sector this would be the quantity of carried passengers, cargo and mail; accrued flying time and aircraft take-offs. The given group of expenditures would also include expenditures on GSM, amortization, flight personnel wages, maintenance and current repair of subsystems, a portion of expenditures on the upkeep of runways and so forth.

The conditionally fixed expenditures, with an increase in the volume of work, either remain constant or rise to a smaller degree than the amount of work and the number of operations performed by the system. They are called conditionally fixed because they do not change until a certain limit is reached for an increase in the amount of traffic. Upon reaching this limit, the expenditures which do not depend upon the amount of traffic jump and then for a certain period again remain fixed (or almost fixed) until a new rise in the amount of traffic causes them to jump again.

Thus, an increase in the number of aircraft take-offs does not change the total expenditures on the amortization of the airfield runway (VPP) and the airport administrative and management expenses. But when the number of aircraft take-offs exceeds airfield capacity and the capability of the airport staff, the VPP must be extended or a new one built, personnel must be increased and so forth. Then the expenditures on BPP amortization and the wages of administrative and management personnel will grow and the expenditures which do not depend upon the amount of traffic as a whole for the airport increase. However, per aircraft take-off these expenditures can be less than before.

4.5.2. Forecasting BTS Operating Expenditures

Depending upon the purpose of the research for the required calculation accuracy BTS operating expenditures can be set with a varying degree of detail. In the general instance their amount is calculated according to the equality

$$\tilde{C}_e = C_\tau(1 + E_n K_{yde})\tau_e, \qquad (4.85)$$

where C_τ--BTS operating costs per unit of time;
E_n--normed sectorial capital investment effectiveness coefficient;
K_{yde}--proportional capital investment;
τ_e--duration of examined BTS operating period.

Generally the forecasting method for BTS operating expenditures is carried out using an aggregate mathematical economics model in which the cost indicators and the basic characteristics of the system and operating stage are analytically interrelated.

Below, from the example of air transport BTS, the method is given for calculating expenditures on their operation per unit of time (for 1 hour of accrued flying time).

Direct operating expenditures (PER). The direct expenditures (see Table 4.7) include the expenditures the amount of which can be directly related to the cost of a system flight hour. These are expenditures on the wages of flight personnel (LPP), fuel and lubricants (GSM), amortization and maintenance of the aviation BTS.

LPP wages. The hourly wages of the LPP with the existing wage system over the entire aircraft life can remain constant, as the amount of the salaries and bonuses with kilometrage wages does not depend upon the aircraft amortization life or the period between repairs.

Proportional LPP wages can be determined when one knows the number of passenger seats or when the number of crew members, the amount of the annual flying hours per crew and the take-off weight of the aircraft are known. However, the number of passenger seats and crew members shows a correlation dependence upon the aircraft take-off weight G_o. As a consequence, LPP wages depend most upon take-off weight and can be calculated from the following equation:

$$G_{1\tau 1pp} = a_{G_o} G_o^{\lambda G_o}. \qquad (4.86)$$

GSM expenditures. In a general form the hourly GSM expenditures are determined in the following manner:

$$G_{1\tau_{gsm}} = Q_a(1+0)P_T, \qquad (4.87)$$

where Q_a--average hourly fuel and lubricant consumption by all aircraft engines in the air, tons;
0--coefficient considering additional GSM consumption in operation of engines on ground;
P_T--price of 1 ton of fuel, rubles.

As is known, the average hourly fuel consumption Q_a changes depending upon the cruising hourly fuel consumption Q_{cr} and the length of the non-stop flight L_{ns}. In turn, the cruising hourly fuel consumption shows a correlation dependence upon aircraft productivity $V_{cr}G_o$, where V_{cr}--cruising speed of flight. Thus, having expressed in the equation (4.87) Q_a by L_{ns} and G_o, we obtain the following type of calculation formula for hourly GSM consumption expenditures:

$$C_{1\tau_{gsm}} = a_{gsm}(V_{cr}G_o)^{\lambda_1}L_{ns}^{\lambda_2}, \qquad (4.88)$$

where $\lambda_2 < 0$.

Amortization. Amortization deductions include outlays on renovation and major overhaul of the airframe with equipment and engines. Hourly renovation deductions are calculated as the quotient of dividing the cost of the functional element minus its residual amount C_{res} by its amortization life τ_i:

$$C_{1\tau}^{p} = \sum_{j=1}^{n} \frac{C_{otr} - C_{res}}{\tau_i}, \qquad (4.89)$$

where C_{otr}--sectorial cost including expenditures on NIOKR and serial production of functional element;
n--number of elements in system. The amount of the residual value C_{res}, depending upon the type of element, varies within the limits of 2-8 percent.

The proportional (per unit of use time) expenditures on major overhaul are determined as the ratio of the cost of the entire quantity of major overhauls over the service life to the amortization life of the functional element.

The cost of a major overhaul on an airframe with equipment can be calculated depending upon the weight of the empty aircraft G_e, the ordinal number of repairs n_{krpl} and the size of the repair inventory N_{rpl}

$$C_{krpl} = a_{kr}G_e^{\lambda_{1n}^{p}}N_{rpl}^{\lambda_{2n}^{p}}n_{krpl}^{\lambda_{3n}^{p}} \qquad (4.90)$$

where $\lambda_{2n}^{p} < 0$.

The cost of an engine major overhaul is calculated depending upon the thrust R_{max} or the power N_{max} for the ordinal number of the overhaul n_{krdv}, and the amount of the repair inventory N_{rdv}:

$$C_{krdv} = a_{kr}R_{max}^{\lambda_{1d}^{p}}\,N_{rdv}^{\lambda_{2d}^{p}}\,n_{krdv}^{\lambda_{3d}^{p}}, \qquad (4.91)$$

where $\lambda_{2d}^{p} < 0$.

Considering that an aircraft can have several engines, the proportional expenditures on major overhauls should be increased by the number of engine mountings m_{dv}.

Expenditures on maintenance. The maintenance process has a cyclical nature (it is repeated after each major overhaul). For this reason aircraft maintenance expenditures related to a flying hour $C_{\tau to}$ can be determined depending upon the length of time between repairs τ_{mr}. In addition it has been established that these expenditures depend upon the take-off weight of the aircraft:

$$C_{\tau to} = \frac{1}{\tau_s} a_{to} G_0^{\lambda_{1to}} \tau_{mr}^{\lambda_{2to}}; \qquad (4.92)$$

where τ_s--amortization life of aircraft.

Table 4.9

1 Класс аэропорта	2 Взлетная масса самолета, т	3 Среднегодовой объем работы аэропорта, тыс. ткм	4 Коэффициент аэропортовых расходов
I	$G_0 > 80$	>100	0,18
II	$80 \geqslant G_0 > 40$	100—30	0,30
III	$40 \geqslant G_0 > 10$	30—10	0,45
IV	$10 \geqslant G_0 > 1,5$	<10	0,6

Key: 1--Airport class; 2--Aircraft take-off weight, tons; 3--Average annual amount of aircraft operations, 1000 ton-km; 4--Airport expenditure coefficient

Indirect operating expenditures (KER). In calculating the cost of an aircraft flight hour, in addition to the direct expenditures, it is also essential to consider the so-called airport (airfield) expenditures which are part of the KER group. At present the airport expenditures are set proportionately to the direct (flight) expenditures or to the aircraft take-off weight. Either method leads to substantial inaccuracies and does not make it possible to optimize the series of aircraft parameters and primarily the aircraft take-off and landing characteristics.

Airport expenditures are significantly influenced by the airport class required for the designed aircraft and the closely related amount of traffic for these airports. In this regard, with aggregated methods for calculating the cost of an aircraft flight hour, the amount of the airport expenditures should be determined proportionally to the direct operating expenditures (the calculation base), but in a differentiated manner for the required airport class or amount of traffic (Table 4.9).

Proportional capital investments. In an economic estimate of the designed aircraft and in optimizing their parameters, it is essential to consider not only the current expenditures but also the one-shot (capital) expenditures. Direct capital investments include investments into the required aircraft and engine fleet and into airport productive capital. Since the traffic volume is measured in ton-km, it is advisable to calculate the proportional capital investments (or proportional productive capital) f_{IW} per 1 ton-km. A transition to proportional capital investments per flying hour $f_{1\tau}$ in this instance is rather simple:

$$f_{1\tau} = f_{1W} W_{1\tau}, \qquad (4.93)$$

where $W_{1\tau}$--hourly productivity of designed aircraft, ton-km,

$$W_{1\tau} = v_c G_{p\ell} K_{\ell}; \tag{4.94}$$

here v_c--average cruising speed, km per hour;
$G_{p\ell}$--payload weight, tons;
K_{ℓ}--load factor.

The proportional capital investments into the aircraft and motor fleet can be calculated proceeding from the forecast amount of aircraft cost C_s and engine cost C_d in their output program N_s and N_d and annual productivity W_{yr}:

$$f_{1W} = \frac{\tilde{C}_s N_s + \tilde{C}_d N_d}{W_{yr}}; \tag{4.95}$$

$$W_{yr} = W_{1\tau} H_{yr}; \quad N_d = 1.15 m_d N_s,$$

where H_{yr}--accrued annual aircraft flying time;
m_d--number of engines on aircraft;
1.15--coefficient considering required engine replacement stock.

Consequently,

$$f_{1W} = \frac{\tilde{C}_c + 1.15 m_d \tilde{C}_d}{W_{1\tau} H_{yr}}. \tag{4.96}$$

The proportional capital investments into airport productive capital can be calculated on the basis of empirical data for the various classes of airports.

CHAPTER 5: METHODS OF DETERMINING THE ECONOMIC EFFECTIVENESS OF LARGE TECHNICAL SYSTEMS

5.1. General Characteristics of Methods

According to the above-accepted classification of effects, it is essential to distinguish the annual effect and the full effect over the entire service life of new equipment. The concept of the full effect can be broadened if one considers not one specimen but rather the entire output of the given model of new equipment or the entire output of equipment needed to implement a certain national economic program.

In the literature, along with the term "full effect," there is also found the concept of the integral effect which is employed in estimating the effect from carrying out integral programs over 10-15 and more years. V. P. Krasovskiy notes that the advantages of integral calculations are that they provide an amount of effect over an extended period of operating the objects or means of new technology considering both the gradual increase in the effect during the production development process as well as its possible decline due to obsolescence. Furthermore the integral effect calculations provide an opportunity to more fully realize a national economic approach to selecting the investment programs, to anticipate possible changes in the market conditions, the consequences of further technical development and the future social conditions as well as caution against hurried and unsound decisions taken solely from the short-term gains of the next few years. The integral effect calculations over an extended period correspond more to the tasks of investment programs than do the annual effect calculations [14, page 19].

We must fully agree with the conclusions of V. P. Krasovskiy. In estimating aviation equipment, an integral approach has long been employed, in truth, under different names such as a comprehensive approach, a dynamic approach and determining the effect over the system's life cycle.

The annual effect can be calculated for a certain year as an average annual effect over the entire service life. If the annual effect is calculated as an average annual effect, then there are no fundamental differences between the methods of calculating the annual effect and the full effect and these differences are purely of a quantitative nature. If the annual effect is calculated for a certain year, then these calculations have a definite static nature in contrast to calculations for the entire system life cycle which are of a dynamic nature. Proceeding from this it is possible to propose two methods for calculating the economic effectiveness of the BTS: dynamic and static.

A dynamic approach makes it possible to consider the dynamics of a phenomenon which arise under the influence of a large number of changing factors and for this reason the results of the dynamic approach are more accurate than the static one. However, the use of a dynamic approach requires a large number of initial data which, depending upon the nature of the calculation, have a forecast or planned nature and encompass a significant period. With a static approach the initial data are more reliable and accurate and for this reason it possesses greater certainty than the dynamic approach. This particular feature of the static model and its simplicity have been responsible for its wide use.

However, the development of methods for planning and forecasting the economic parameters of technical systems and the introduction of program planning and forecasting methods have set a trend to an ever-greater transition from the static to the dynamic methods of calculating system economic effectiveness. Evermore elements of the dynamic method are being incorporated in the static method. Subsequently we will be talking about the dynamic method as the most general calculation method. Along with this we can recommend an approximate calculation method which combines individual elements of the static approach to elements of the dynamic method.

The use of one of the two recommended methods--dynamic and approximate--depends upon the nature of the problem to be solved and the amount of initial information. Global problems, for example, the involving of the economically sound periods for the change of BTS generations or determining economic effectiveness of replacing a functioning system by a new one and so forth, can be most successfully solved by the dynamic method. The local (particular) problems which can be solved with a minimum of initial data, for example, the choice of optimum parameters for a system element in terms of an economic criterion with the set basic system parameters or a feasibility study for an element to be used in a given system and so forth, can be successfully solved using the approximate method.

5.2. The Dynamic Method for Determining National Economic Effectiveness of the BTS

An estimate of BTS economic effectiveness assumes the determining of two effectiveness criteria; the national economic and the cost accounting effect. The solution to the problems of determining BTS economic effectiveness using these criteria assumes the constructing of mathematical economics models which show the dynamics of capital investments and current expenditures on NIOKR as well as in serial production and operation over the entire system life cycle period.

The dynamic calculation method is based upon the constructing of a dynamic model for the BTS economic effectiveness criterion, that is, a model of expenditures on the development, manufacture and operation of the BTS. In constructing this model, major difficulties arise in the mathematical description of capital investment and current expenditures dynamics over the years. A dynamic expenditure model should be based upon the model for the production and reproduction of the entire BTS fleet over a significant period of time. It is very complicated to construct such a model, as hundreds and even thousands of articles are in operation. The BTS fleet circulates from the serial production plant to operation, then to repairs and back into operation. It is complex to construct a model for the production, reproduction and operation of the BTS fleet in a natural scale, however, it is significantly more complicated to construct an expenditure model, as the articles manufactured at different times differ both in terms of expenditures and in the calendar time of their making.

The acquired experience in constructing dynamic models of the BTS economic effectiveness criteria makes it possible to formulate the basic provisions for the elaboration of such models. A dynamic model for the BTS economic effectiveness criterion represents a system of time-coordinated particular dynamic models for the production, reproduction and operation of the BTS (a model of the natural indicators) as well as dynamic models of expenditures made over the BTS life cycle stages (the model of cost indicators).

For the practical use of the designated dynamic model, of crucial significance is the aggregate of initial data used as the basis for extimating BTS economic effectiveness. The most important element in the initial base is the national economic demand which the designated BTS should satisfy. The national economic demand is characterized by a range of qualitative and quantitative parameters which determine the amount of work produced by the given system over the period of operating the system T_0. The second element in the initial base for constructing the model is the range of parameters characterizing the quality of the designated system variations.

The amount of the system's work over time during T_0 years and the values of the system parameters, particularly those which characterize its productivity and operational effectiveness, make it possible to set the required number (fleet) of system elements. Let us recall that the elements of an aviation system are the aircraft and its structural parts: the airframe, power unit, flight equipment and the ground support facilities.

The third initial element for constructing the model is the available material and human resources needed for the creation, production and functioning of the BTS. First of all it is a question of production capacity possessed by the scientific research, prototype design and industrial enterprises which create and manufacture the various system elements as well as the organizations which operate the BTS. A comparison of the demand for system elements with the production capacity of the corresponding organizations and enterprises makes it possible to disclose the lacking production capacity and the amount of required capital investments and human resources.

The calculated demand for the BTS serves as the basis for calculating the output program for the BTS elements at the corresponding enterprises, proceeding from the existing production capacity and capability for building new capacity. Analogous balance calculations are made for the repair enterprises. This makes it possible to determine the time parameters of the system life cycle:

1) The start of capital construction in the various spheres such as scientific research institutes, design organizations, industry and in operation;

2) The duration of capital construction in the corresponding sphere;

3) The beginning of carrying out the scientific research and prototype design work (NIOKR) for the various system elements;

4) The duration of carrying out the NIOKR for the various system elements;

5) The start of manufacturing the various system elements and the duration of the manufacturing period.

All the time parameters should be interrelated and coordinated and precisely in this case do they characterize the system life cycle.

The presence of technical parameters for the designated system elements, the quantitative data on the demand for them, the time characteristics of the system life cycle and data on the availability of production resources and the completion of new resources make it possible to determine the system's cost parameters: the cost of the NIOKR for the various system elements, the cost of manufacturing the system and the cost of operating the system.

Thus, the dynamic model of the BTS economic effectiveness criterion includes the following partial models: the model of the specific BTS effect; the model for the quality parameters of the system elements; the model for the production and operational dynamics of the BTS; the expenditure models in the stages of the system life cycle.

The method of constructing the particular models and synthesizing them into a general effectiveness criterion model is best examined from the example of a certain specific large technical system. An air transport system can serve as an example of a large technical system and for this reason let us examine the method of constructing a dynamic model for the BTS economic effectiveness criterion using the example of an air transport system (ATS).

The national economic effectiveness of an ATS is calculated using the criterion (3.24). The annual effect from introducing the new system is determined by (3.27) and the full effect by (3.28). In a dynamic calculation method, a dynamic model of the ATS effectiveness criterion is employed and this includes the particular models listed above.

Model of the specific effect of the ATS. The ATS variations are compared with a constant fixed amount of the system's specific or telic effect. The specific effect is the volume of air transport work B_Σ which the system carries out over the period of time T_o $[T_{so}, T_{eo}]$. The volume of air cargo shipments is measured in ton-km and passenger traffic in passenger-km while the total volume of cargo and passenger shipments is given in reduced ton-km which are determined by adding the cargo and passenger turnover multiplied by a coefficient of 0.09 indicating how much a passenger with a free baggage allowance weighs in tons (90 kg).

The system's volume of traffic is determined by the assumed spheres of its use, that is, by the transport network which is characterized by the aggregate of air routes set by the length (distance) $L = \{\ell\}$ and by the volume of air shipments $G(\ell, t)$ in reduced ton-km for each distance ℓ from the designated aggregate $\{\ell\}$ for each year t of the period T_o.

Model of quality parameters for the ATS elements. As was pointed out above, the system's elements are characterized by a large number of quality parameters, including: functional, technical and economic. In an economic estimate of a system, one considers not the entire range of parameters but rather the basic ones. For example, for an aircraft which is the central element of the system, the basic initial parameters for the calculation are the following:

1) The given class of aircraft (an aircraft for local lines or trunk lines, passenger or cargo, and so forth);

2) Maximum aircraft payload G_{com}^{max};

3) Planned aircraft load factor K_z;

4) Calculated range $\ell_{ca\ell}$ or the maximum non-stop flight range with the maximum payload;

5) Change in the payload $G_{com\ell}$ and consumed fuel $G_{t\ell}$ depending upon the flight distance (Fig. 5.1) and the change in the cruising speed $V_{cr\ell}$ depending upon the length of flight ℓ;

6) The average annual accrued flying time per aircraft T_g;

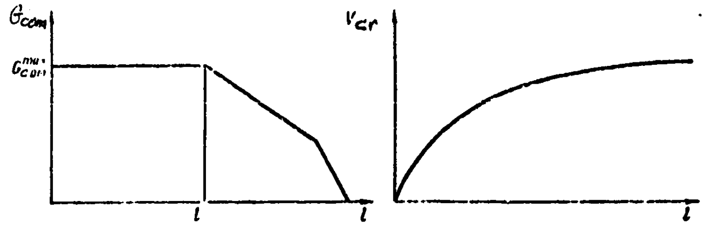
7) Take-off weight G_o, weight of empty aircraft G_{em} and airframe weight G_{af};

8) Number of engines on aircraft N_d;

9) General technical life of aircraft airframe $\tau_{s\Sigma}$;

10) Service life between repairs for airframe τ_s;

11) Number of airframe major overhauls n_{rs}.



a) b)

Fig. 5.1. Dependence of payload (a) and cruising speed (b) on flight range

The basic initial parameters for an engine are:

1) Thrust R_o, R_{cr} and specific fuel consumption C_{Ro}, C_{Rcr} (under take-off and cruising conditions);

2) Operating time between repairs τ_d and general technical life $\tau_{d\Sigma}$ of engines in operation;

3) Number of engine major overhauls n_{rd}.

Model of production and operation dynamics of ATS.[1] For normal operation of an ATS, it is essential to have a certain number of its elements, that is, aircraft, engines and so forth. Let us examine the method of determining the required number of aircraft and engines.

The number of aircraft needed to handle the set volume of air shipments in year t of operations, N_{sot} is calculated from the following formula:

$$N_{sot} = \sum_{l} N_{sot} = \sum_{l} \frac{G_{l,t}}{B_{rl}}, \tag{5.1}$$

where $N_{sot,\ \ell}$--number of aircraft required to handle set volume of air shipments over distance ℓ in year t;

$B_{g\ell}$--annual productivity of one aircraft for distance ℓ.

If it is considered that annual aircraft productivity is:

$$B_{g\ell} = G_{com\ell} V_{cr\ell} T_g K_z,$$

then (5.1) assumes the form

$$N_{sot} = \sum_{\ell} \frac{G_{\ell,\ t}}{G_{com\ell}} V_{cr\ell} T_g K_z. \tag{5.2}$$

The following stage is to calculate the demand for the delivery of new aircraft for operation in such a quantity as to ensure the functioning of the aircraft fleet. For this it is essential to consider not only the growth of the fleet but also the taking of a portion of aircraft out of service at the end of their lives. Let us designate the amount of the demand of operation for receiving new aircraft by N_{soreqt}. The nature of the behavior of the amount N_{soreqt} over the designated period can be rather complex. At the same time, serial production is adapted only to a steady production pace. The problem arises of optimizing the relation between the demand for the ATS elements with their production programs.

Thus it is essential to determine programs for the serial production plants N_m^{max} which would meet the operational requirements for the ATS elements considering reproduction in operation, being at the same time stable programs.

Before examining the calculation algorithm for production programs, let us describe the models for the production and operation of the aircraft fleet and then, using these models, let us construct the calculation algorithm. Let us introduce a concept for the increase in the aircraft fleet ΔN_{sot} which represents the increase in the number of aircraft in operation (increase in the fleet) during year t.

---

[1] Model was worked out jointly with Engr V. M. Dikushin.

The values of the amount ΔN_{sot} can be determined by calculating the difference between the values of the current N_{sot} and previous $N_{so(t-1)}$ years, with the exception of the first year of operation for which the value ΔN_{sot} coincides with the fleet size of the first year:

$$\Delta N_{sot} = \begin{cases} N_{sot} \quad \text{with} \quad t = T_{so} + 1; \\ N_{sot} - N_{so(t-1)} \quad \text{with} \quad T_{so} + 1 < t \leqslant T_{eo}, \end{cases} \qquad (5.3)$$

where T_{so}, T_{eo}--beginning and end of operational period of aircraft fleet, years.

The amount ΔN_{sot} indicates how many aircraft must be produced in year t proceeding from the needs of increasing the fleet (not counting the reproduction of the aircraft fleet). Let us designate by $N_{s(\tau)\Sigma}$ the number of aircraft required for creating the fleet of the given type of aircraft for the period $[T_{so} + 1, T_{eo}]$ not counting the aircraft which have gone for reproduction:

$$N_{s(\tau)\Sigma} = N_{s\Sigma} + N_{wo\ s\Sigma}, \qquad (5.4)$$

where $N_{s\Sigma}$--number of aircraft required for the program of creating the given type of aircraft fleet;
$N_{wos\Sigma}$--the number of written off (taken out of operation) aircraft of the given type over the period T_o.

We would point out that the values of the amounts $N_{s\Sigma}$ and $N_{wos\Sigma}$ do not depend upon what production dynamics and annual program we adopt. The nature of consumption, that is, of operation which is determined by the amount ΔN_{sot}, remains constant. With the various production programs and models, only the volume of the fixed capital of the operating organizations changes. As a consequence of this, the number of aircraft can be calculated from the production model constructed under the assumption that the output dynamics corresponds to the nature of the fleet's growth in operation and the ensuing dynamics of fleet reproduction.

Such a production model N'_{smt} will reflect the behavior of the amount N_{soreqt}. It can be used as a rough approximation for calculating production programs. The amount of N'_{smt} is determined from (5.4):

$$N'^{*}_{smt} = \begin{cases} N_{sot} \quad & \text{with} \quad T_{so} < t < T_{so} + \tau_{s\Sigma}/T_g, \\ N_{sot} + N'_{sm(t-\tau_{s\Sigma}/T_g)} \quad & \text{with} \quad T_{no} + \tau_{s\Sigma}/T_g < t < T_{eo}, \end{cases}$$

where the term $N'_{sm(t-\tau_{s\Sigma}/T_g)}$ considers the reproduction volume of aircraft put into operation $(\tau_{s\Sigma}/T_g)$ years ago.

---

*The prime mark in designating the amount N'_{smt} indicates that the given formula is not a real production model but is used solely to calculate the demand for new aircraft in operation.

The total number of aircraft N_s needed to carry out the set volume of air shipments for the period $[T_{so}+1,\ T_{eo}]$ is determined from the formula

$$N_{s\Sigma} = \sum_{t=T_{so}+1}^{T_{eo}} N'_{smt}, \tag{5.5}$$

while that portion of the aircraft which has worked its life over the designated period $N_{swo\Sigma}$ is calculated by a separate totaling of the term $N'_{sm}(t-\tau_{s\Sigma}/T_g)$ using (5.5)

$$N_{swo\Sigma} = \sum_{t=T_{so}+1+(\tau_{s\Sigma}/T_g)}^{T_{eo}} N'_{sm}(t-\tau_{s\Sigma}/T_g).$$

Analogous calculations can be made for the engine fleet.

Let us designate by $N_{d(\tau)}$ the number of engines required for the initial making up of the aircraft fleet $N_{s(\tau)}$:

$$N_{d(\tau)} = N_{s(\tau)\Sigma}N_d. \tag{5.6}$$

The amount of $N_{d(\tau)\Sigma}$ determines the amount of capital investments into the engine fleet. All the engines above this amount go to reproduce [replace] the initial ones. The total demand for engines over the designated period $N_{d\Sigma}$ can be determined by the formula

$$N_{d\Sigma} = N_{s(\tau)\Sigma}\frac{T_{s\Sigma}}{\tau_{d\Sigma}}, \tag{5.7}$$

where the amount $\tau_{s\Sigma}/\tau_{d\Sigma}$ represents the ratio of the general technical life of the airframe $\tau_{s\Sigma}$ and the engine $\tau_{d\Sigma}$ determining the number of engines per mounting and required for the normal operation of the aircraft until the end of its full operating life.

The required number of operations in operation (on the aircraft or in replacement stock) during year t is determined by the formula

$$N_{dot} = N_{sot}n_d(1+K_o), \tag{5.8}$$

where K_o--coefficient considering that a portion of the engines is under repair, in storage, en route.

The value N_{dot} is analogous to N_{sot} and is related to it by a directly proportional dependence. As for the aircraft, let us determine the increased demand for engines in operation N_{dot} which is related to an increase in the volume of air shipments carried out by the given fleet:

$$\Delta N_{dot} = \begin{cases} N_{dot} & \text{with } t = T_{so}+1; \\ N_{dot} - N_{do(t-1)} & \text{with } T_{so}+1 < t \leqslant T_{eo}; \end{cases} \qquad (5.9)$$

$$\Delta N_{dot} \ngtr 0 \qquad \text{with } T_{so} < t \leqslant T_{eo}.$$

Let us move on to determining the serial production programs for aircraft and engines. Let us assume that production has the nature corresponding to models 1 and 3 (Fig. 5.2). Model 3 corresponds to the production dynamics of a new aircraft and new engine $K_d = 0$), while model 1 is for an already existing aircraft and engine ($K_d = 1$).

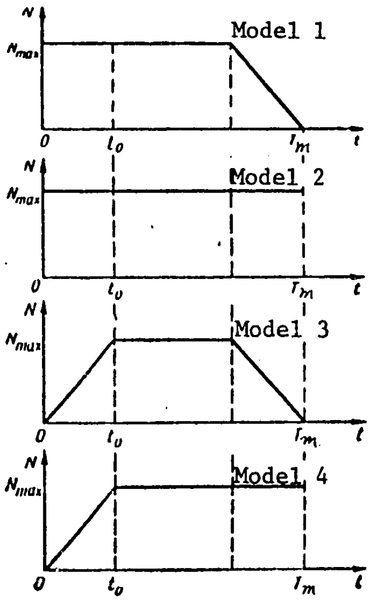


Fig. 5.2. Engine output models

Let us assume that we know the range of discret values for the amounts of the programs $N_{s,\ dm}$ determined proceeding from possible production with the varying degree of utilizing production capacity for the given production program. Let us designate by $N_{s,\ dmo}$ the initial value of the range of programs, and by $\Delta_{s,\ dmk}$--the difference between the sequential values of the production programs from the set range:

$$\Delta_{s,\ dmk} = N_{s,\ dm(k+1)} - N_{s,\ dmk},\ k = 0.1,\ \ldots,\ K-1,$$

where K--number of divisions in set range of programs.

Let us also assume that we know a certain production program from the number set $N_{s,dmk}$ which knowingly satisfies the operational requirements for any year t. The corresponding production duration $T_{s,\ dm}$ is calculated from the formulas

$$T_{s,\ dm} = \frac{N_{s,\ d\Sigma}}{N_{s,\ dmk}} + T_{s,\ do} \quad \text{for} \quad K_d = 0 \qquad (5.10)$$

or

$$T_{dm} = \frac{N_{d\Sigma}}{N_{dmk}} + \frac{1}{2} T_{do} \quad \text{for} \quad K_d = 1. \qquad (5.11)$$

Then, in accord with the adopted production graphs (see Fig. 5.2), the aircraft and engine production models for $K_d = 0$ can be written in the following form:

$$N_{s,\ dmt} = \begin{cases} \dfrac{N_{s,\ dmk}}{T_{s,\ do}+1}\, t \quad \text{with} \quad 1 \leq t \leq T_{s,\ do}; \\ N_{s,mk} \quad \text{with} \quad T_{s,\ do} < t \leq T_{s,\ dm} - T_{s,\ do}; \\ \dfrac{N_{s,\ dmk}}{T_{s,\ do}+1}\,(T_{s,\ dm}+1-t); \\ \text{with} \quad T_{s,\ dm} - T_{s,\ do} < t \leq T_{s,\ dm}, \end{cases} \tag{5.12}$$

and for an existing aircraft and engine, that is $K_d = 1$, formula (5.12) assumes the form

$$N_{s,\ dmt} = \begin{cases} N_{s,\ dmk} \quad \text{with} \quad 1 \leq t \leq T_{s,\ dm} - T_{s,\ do} \\ \dfrac{N_{d,\ mk}}{T_{s,\ do}+1}\,(T_{dm}+1-t) \\ \text{with} \quad T_{s,\ dm} - T_{s,\ do} < t \leq T_{s,\ dm}. \end{cases} \tag{5.13}$$

Formulas (5.12) and (5.13) are obtained by dividing the production dynamic graphs into stages with uniform processes. Thus, in formula (5.12), the production period $T_{s,\ dm}$ is broken up into three intervals: 1) starting up of production--the period of developing serial production; 2) serially developed aircraft output; 3) curtailment of production. For each of these intervals there is a formula corresponding to the process occurring in it.

In the first interval, the ratio $\frac{N_{s,\ dmk}}{T_{s,\ do}+1}$ indicates the annual production increase. In multiplying the amount of the annual production increase by the number of the production year, we obtain the number of produced aircraft or engines in the current year within the limits of the serial production period $1 \leq t < T_{s,\ do}$. In the second interval there is the even output of aircraft and an even $N_{s,\ dmk}$. The formula for the third interval, as is seen fromthe graphs of Fig. 5.2, reflects the process of production curtailment going on in a procedure inverse to the process of developing production.

Let us set the following problem. It is necessary to determine the value for the annual production program $N_{s,\ dm}$ which, with the given production model, would meet the operational requirements in any year t and be the minimum possible number of all the set values of $N_{s,\ dmk}$. Let us designate the sought value by $N^{min}_{s,\ dm}$. Thus, we are searching for the lower level of the production program which would meet production requirements. For calculating $N^{min}_{s,\ dm}$ let us adopt the following altorithm.

1. Let us appropriate the values: $t = 1$, $k = 0$.
2. Let us appropriate to the value $N^{min}_{s,\ dm}$ the value of the greatest values of the amount $N_{s,\ dmk}$ or $N_{s,\ dot}$:

$$N^{\min}_{s,\ dm} = \max\{N_{s,\ dmk},\ \Delta N_{s,\ dot}\},$$

where $\Delta N_{s,\ dot}$--increase of fleet in operation during year t. In the first step $\Delta N_{s,\ dot} = \Delta N_{s,\ de}$.

3. On the basis of the obtained values for the amount $N^{\min}_{s,\ dm}$, with the various output models, let us calculate the production cynamics models sequentially for (5.4)-(5.7), (5.10)-(5.13), that is, we will obtain the values $N'_{s,\ dmt}$, $N_{s\Sigma}$, $N_{d(\tau)\Sigma}$, $N_{d\Sigma}$, T_{ms}, T_{md}, N_{smt}, N_{dmt}.

4. For each year t let us calculate the difference of the sums ω_m of the following type:

$$\omega_m = \sum_1^t N_{s,\ dmt} - \sum_{t=T_{so}+1}^{t} [\Delta N_{s,\ dot} + N'_{s,\ dm}(t-\tau_{s,\ d\Sigma}/T_g)].$$

5. Let us verify the sign of the amount ω_m:

a) If ω_m is negative in any t, let us increase the value k by one, that is, $k = k+1$.

$$N^{\min}_{s,\ dm} = N^{\min}_{s,\ dm} + \Delta_{s,\ dmk}$$

and return to step 2;

b) If all the possible values for the production programs ($k = K$) are exhausted, additional capacity is required.

c) If the amount ω_m is non-negative for all the values of t, then we have completed the calculation, and if at a certain point of t the amount of ω_m is negative, then we increase k by one and repeat the calculation starting at step two.

Thus, as a result of the designated procedure we obtain the minimum acceptable level of annual output which can be used as a constraint in the general problem of optimizing the ATS parameters in terms of the economic effectiveness criterion. For completing the description of the fleet dynamics models it is necessary to describe the dynamics of aircraft and engine repairs.

The number of aircraft repairs during year t is determined from the formula

$$N_{rst} = \begin{cases} 0 \text{ with } T_{so} < t \leqslant T_{so} + \dfrac{T_s}{T_g}; \\ N_{so}(t-\tau_s/T_g) + N_{sm}(t-\tau_s/T_g) \text{ with } T_{so} + \tau_s/T_g < t \leqslant T_{so} + \tau_{s\Sigma}/T_g; \\ N_{rs}(t-\tau_s/T_g) + N_{sm}(t-\tau_c/T_g) - N_{sm}(t-\tau_{s\Sigma}/T_g) \\ \text{with } T_{no} + \tau_{s\Sigma}/T_g < t \leqslant T_{eo}, \end{cases} \tag{5.14}$$

where τ_s/T_g--the calendar time for the operation of the aircraft between two major overhauls;

$\tau_{s\Sigma}/T_g$--service life of one aircraft.

In the event that in calculating the ratios τ_s/T_g and $\tau_{s\Sigma}/T_g$ we obtain fractions, it is essential to take the whole number closest to the obtained result.

The number of engine overhauls in year t is determined from the formula

$$N_{rdt} = \begin{cases} 0 \text{ with } T_{so} < t \leqslant T_{so} + \dfrac{\tau_d}{T_g}; \\ N_{dm}(t-\tau_d/T_g) + N_{rd}(t-\tau_d/T_g) \text{ with } N_{so} + \tau_d/T_g < t \leqslant T_{no} + \tau_{d\Sigma}/T_g; \\ N_{dm}(t-\tau_d/T_g) + N_{rd}(t-\tau_d/T_g) - N_{dm}(t-\tau_{d\Sigma}/T_g) \\ \text{with } T_{so} + \tau_{d\Sigma}/T_g < t \leqslant T_{eo}. \end{cases} \qquad (5.15)$$

The general principle for construction (5.14) and (5.15) consists in dividing the operational interval into intervals with uniform processes. Thus, in (5.14) in the time interval $T_{so} < t \leqslant T_{so} + \tau_s/T_g$, the number of overhauls $N_{rst} = 0$, since none of the aircraft had yet operated the amount of time between repairs.

In the interval $T_{so} + \tau_s/T_g < t \leqslant T_{so} + \tau_{s\Sigma}/T_g$, the aircraft manufactured in τ_s/T_g years before the current moment of time T begin to go into repair in the number $N_{sm}(t-\tau_s/T_g)$, and a second overhaul during the following τ_s/T_g years $N_{rs}(t-\tau_s/T_g)$. Obviously during the first years (up to $T_{so} + 2\tau_s/T_g$) of this interval, the amount $N_{rs}(t-\tau_{s\Sigma}/T_g)$ will equal zero, in repeating the values of the amount N_{rst} in the first interval. As yet there has been no writing off of the aircraft.

In the third interval $T_{so} + \tau_{s\Sigma}/T_g < t \leqslant T_{eo}$, a portion of the aircraft begins to be written off instead of undergoing repair, and for this reason from the amount of repairs $N_{rs}(t-\tau_s/T_g)$ it is essential to subtract the number of aircraft to be written off produced $\tau_{s\Sigma}/T_g$ years ago from the current moment of $N_{sm}(t-\tau_{s\Sigma}/T_g)$.

Expenditure models for stages of system life cycle. The obtained models of the physical indicators make it possible to move on to constructing expenditure models for the ATS life cycle stages. Expenditures in the various life cycle stages vary and they include current and capital expenditures. Initially let us examine a model of current expenditures for the entire life cycle of the system T_Σ. The current expenditures considered in the criterion of BTS national economic effectiveness would include current expenditures on NIOKR for each element C_{jniokr}, current expenditures on manufacturing the system elements C_{jm} and current expenditures on operating the system elements C_{jo}.

Current expenditures on NIOKR for each system element C_{jniokr} include the expenditures on NIR C_{jnir} and on OKR C_{jokr}. Current expenditures on NIR can be determined proceeding from the dependence of (4.35):

$$C_{jnir} = C_{jokr} K_{jnir}, \qquad (5.16)$$

where K_{jnir}--coefficient characterizing the ratio between current expenditures on NIR and OKR.

Thus, the OKR expenditures are the basis for determining the NIR expenditures. The method of forecasting OKR expenditures has been examined in Chapter 4. If, for example, we take the central element of the system, the aircraft, then the expenditures on creating it can be calculated using an empirical equation as a function of the parameters X_j (maximum take-off weight, maximum speed and so forth):

$$C_{jokr} = b \prod_{j=1}^{m} X_j^{b_j}, \tag{5.17}$$

where b and b_j--empirical coefficients.

Models of the type (5.17) make it possible to forecast the total OKR expenditures. At the same time the criterion considers the annual current OKR expenditures. The annual OKR expenditures C_{jokrt} can be determined on the basis of the total expenditures. In knowing the total expenditures C_{jokr} and in knowing the standard distribution of expenditures over the years η_{okrt} indicating what percent of the total expenditures is spent in year t, it is possible to estimate the annual current OKR expenditures for element j:

$$C_{jokrt} = C_{jokr}\eta_{jokrt}.$$

Above, in Chapter 4, the basic formative patterns and methods of forecasting BTS element costs were established. Three stages were examined in the process of serial production and conditional limits were set between the individual stages of the process. The limit between the stage of developed serial production was called the beginning of the period of serially developed production. If one measures the time from the beginning of production in quarters, then the start of the period of serially developed production is characterized by a certain quarter with the start of production t_0. The average cost of the articles manufactured in the given quarter will characterize the cost of the serially produced article C_0.

The cost of a serially produced product characterizes the cost level at a certain moment of time. For describing not only the level but also cost dynamics, it is essential to elaborate a model which would describe cost dynamics from the production parameters. Above it was pointed out that for characterizing production factors which influence product cost dynamics, the following indicators are used: the number of manufactured articles since the start of production N, the calendar time since the start of production t and the daily product output q. The product cost model reflecting the cost level and dynamics can be represented in the following form:

$$C_N = C_0K_N; \tag{5.18}$$

$$C_{N,\ t} = C_0K_{N,\ t}; \tag{5.19}$$

$$C_{t,\ q} = C_0K_{t,\ q}. \tag{5.20}$$

where K_N, $K_{N,\ t}$ and $K_{t,\ q}$--coefficients of cost dynamics depending upon the indicators N, t and q.

The cost models for the system elements are constructed on the basis of the above-given concepts. Let us examine these models from the example of two ATS elements, the aircraft and the engines. The aircraft cost model is constructed in the following manner. The cost of the serially produced aircraft C_{so} is determined as a function of the aircraft functional parameters, that is, weight, speed and so forth. On the basis of the amount of C_{so}, the average cost of one aircraft in year t is determined or C_{s1mt}. The total aircraft production cost in year t equals the product of the cost of one aircraft C_{s1mt} by the annual output N_{smt}.

The cost of the engines in year t C_{dmt} is determined analogously as the product of the cost of one engine in year t C_{d1mt} by the number of engines manufactured in the same year N_{dmt}. The cost of a serially produced engine C_{do} is determined from empirical formulas depending upon its parameters such as thrust, specific fuel consumption and so forth. The cost of one article (of an aircraft C_{s1mt}, of an engine C_{d1mt}) is determined from formulas (5.18)-(5.20) on the basis of the cost of a serially produced article and the coefficients K_N, $K_{N,\ t}$ and $K_{t,\ q}$.

The cost dynamics coefficient K_N is determined from the empirical equation $K_N = b_1 N^{b_2}$ where N--total number of articles (aircraft, engines) manufactured from start of production to the middle of year t which can be designated by $N_{m\Sigma t}$; b_1 and b_2--parameters of the equation

$$N_{m\Sigma t} = \sum_{t=T_{sm}+1}^{t-1} N_{mt} + \frac{N_{mt}}{2}\ . \tag{5.21}$$

In the given formula ones takes one-half the amount of output in year t in order to obtain the average cost of an article during year t. The cost dynamics coefficient $K_{N,\ t}$ is determined from the empirical equation

$$K_{N,\ t} = b_1'[N_{m\Sigma t}]^{-b_2'}\, t^{-b_3},$$

where b_1', b_2' and b_3--empirical equation parameters.

As for the coefficient $K_{t,\ q}$, the deriving of the formula to calculate it needs certain clarifications. The production indicators characterizing in the given case the cost level and dynamics are the amount of the daily output q and the time since the start of production t. The influence of the q and t indicators on cost dynamics and level can be represented in the form shown in Fig. 5.3. Let us assume that three plants produce the same articles (1, 2 and 3) but with different daily outputs (q', q'' and q'''). Let us assume that the daily product outputs in the first quarter (q_1', q_1'' and q_1''') and at the beginning of the serial production period (q_0', q_0'' and q_0''') are characterized by the following ratios:

$$q_1' = q_1''; \ q_1''' > q_1';$$
$$q_0' > q_0''; \ q_0'' = q_0'''.$$

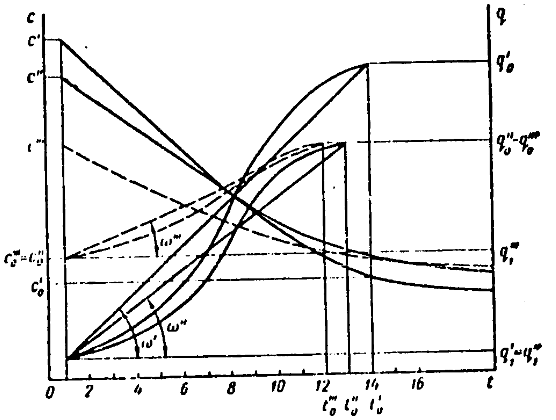


Fig. 5.3. Diagram of influence of t and q indicators on product costs C

The amounts t_0', t_0'' and t_0''' euqal, respectively, 14, 13 and 12 quarters. The designated articles 1 and 2 differ in the amounts of daily output at the beginning of the serial production period and, consequently, during this period their cost levels should differ as is shown in Fig. 5.3. For the articles 2 and 3, $q_0'' = q_0'''$, and hence $C_0'' = C_0'''$. Since $q_0' > q_0''$, then $C_0' < C_0''$. Product cost dynamics are influenced by the amounts of daily output at the beginning of the serial production period q_1 and in the first quarter q_0 by the duration of the production development period t_0 as well as by the nature of the increase in daily output over the quarters.

Research on daily output dynamics over time and its influence on expenditure dynamics has been carried out in Chapter 4. If one somewhat simplifies the problem and accepts that the dependence q = f(t) is of a linear nature, then the cost reduction curves C', C'' and C''' can be presented in the form

$$C' = b_1' t^{-b_3'};\ C'' = b_1'' t^{-b_3''};\ C''' = b_1''' t^{-b_3'''}. \tag{5.22}$$

In these equations, the exponent b characterizes the cost reduction dynamics and depends upon q_1, q_0 and t_0. The increase rate of daily output is characterized by the index for the tangent of angle ω:

$$\operatorname{tg} \omega = \frac{q_0 - q_1}{t_0}. \tag{5.23}$$

The problem comes down to determining $b_3 = f(\operatorname{tg} \omega)$. The dependence of the index b_3 upon the amount of tg ω is approximated by the following curve:

$$b_3 = b_4 (\operatorname{tg} \omega)^{b_5}. \tag{5.24}$$

Having determined tg ω from (5.23) and then from equation (5.24) the value of the index b , it is possible to establish the engine cost reduction dynamics using the formula (5.22).

Now let us determine a model for current system operating expenditures. The operating costs C_{ot} are formed from the operating costs of the basic elements, that is, the aircraft, the engines and airport expenses:

$$C_{ot} = C_{samt} + C_{stt} + C_{w\ell ptt} + C_{damt} + C_{dtt} + P_{ft} + C_{tht} + C_{apt}, \tag{5.25}$$

where C_{samt}, C_{damt}--amortization deductions for aircraft (engines) in year t;
C_{stt}, C_{dtt}--expenditures on maintenance of aircraft (engines) during year t;
$C_{w\ell ptt}$--wages of flight personnel in year t;
P_{ft}--expenditures on fuel in year t;
C_{otht}--other expenditures in operation;
C_{apt}--airport expenses in year t.

Amortization deductions in year t include expenditures on the major overhaul of the aircraft and engine and renovation deductions and are calculated by the formulas:

$$C_{samt} = P_{srt} + \frac{\overline{P}_s T_g}{\tau_{s\Sigma}} N_{sot}; \tag{5.26}$$

$$C_{damt} = P_{drt} + \frac{\overline{P}_d T_g}{\tau_{d\Sigma}} N_{dot}, \tag{5.27}$$

where P_{srt}, P_{drt}--expenditures on major overhaul of aircraft (engines) in year t;
$\overline{P}_s$, $\overline{P}_d$--average price of aircraft (engine) for entire manufacturing period;
$\tau_{s\Sigma}$, $\tau_{d\Sigma}$--General technical life of aircraft (engine).

Expenditures on major overhauls during year t of operation are determined from the formula

$$P_{s,\ drt} = \gamma_{s,\ dr} \overline{P}_{s,\ d} N_{s,\ drt}, \tag{5.28}$$

where $\gamma_{s,\ dr}$--the cost of one major overhaul of an aircraft (engine) in fractions of its price.

The average price of the aircraft (engine) is:

$$\overline{P}_{s,\ d} = \overline{C}_{s,\ d} + R_n F_{s,\ d}, \tag{5.29}$$

where $\overline{C}_{s,\ d}$--average cost of one aircraft (engine) over period of serial production;
$F_{s,\ d}$--proportional capital intensiveness in serial production of one aircraft (engine);
R_n--normed profitability.

As is seen from (5.29), the average price of an article is set on the basis of the average cost over the period of serial production. This cost has been named the average aggregate cost. For example, if over a certain period of time N articles was manufactured, then a distinction is drawn between the average aggregate cost of one article $\overline{C}_n$ from batch N in contrast to the so-called individual cost of N article C_N. Depending upon what indicators are used to describe the conditions of serial production N; N, t or t, q, it is possible to distinguish also the types of average aggregate costs $\overline{C}_N$, $\overline{C}_{N,\ t}$ and $\overline{C}_{t,\ q}$. The amount of the average aggregate cost is determined by formulas analogous to (5.18)-(5.20):

$$\overline{C}_N = \overline{C}_0 \overline{K}_N;$$
$$\overline{C}_{N,\ t} = \overline{C}_0 \overline{K}_{N,\ t};$$
$$\overline{C}_{t,\ q} = \overline{C}_0 \overline{K}_{t,\ q},$$

where $\overline{C}_0$--average aggregate cost of serially produced article;
$\overline{K}_N$, $\overline{K}_{N,\ t}$ and $\overline{K}_{t,\ q}$--dynamic coefficients for average aggregate cost depending upon indicators N; N, t and t, q.

The values of the amounts $\overline{C}_o$, $\overline{K}_N$ and $\overline{K}_{N,\ t}$ are determined by the method analogous to the method of calculating the amounts C_o, K_N and $K_{N,\ t}$ as shown above. The coefficient $\overline{K}_{t,\ q}$ is calculated in the following manner:

$$\overline{K}_{t,\,q}=\frac{\int\limits_0^{T_m} N_{d_m} C_t\,dt}{\int\limits_0^{T_m} N_{d_m} t\,dt},$$

where C_t--the cost at moment of time t; determined by the cost of the serially produced product:

$$C_t = C_0\left(\frac{t}{t_0}\right)^{-b_3}.$$

After integration of the expression $\overline{K}_{t,\ q}$ and certain transformations, we obtain, for example, for model 2 (Fig. 5.2):

$$\overline{K}_{t,\,q}=\frac{2\left[(2-b_3)\,t_0T_m^{1-b_3}-t_0^{2-b_3}\right]}{t_0^{1-b_3}(1-b_3)(2-b_3)(2T_m-t_0)}.$$

The amount of current expenditures on aircraft maintenance in year t is determined from the formula

$$C_{stt} = T_g\,\frac{N_{sot}}{N_{s\Sigma}}\sum_{\ell} C_{sth\ell}\,N_{so\ell} \qquad (5.30)$$

where C_{sth} --hourly expenditures on aircraft maintenance for range ℓ;

$\frac{N_{sot}}{N_{s\Sigma}}$ --coefficient for equal distribution of aircraft in operation during year t;

when the entire fleet is in operation, this coefficient equals one. The amount $C_{sth\ell}$ can be calculated from an empirical formula of the type

$$C_{sth\ell} = \left(a_1+\frac{a_2\bar{F}_s}{d_1}\right)\left(\frac{1}{t_p}+a_3\right)+\left(a_4+\frac{G_{en}}{d_2}+a_5+\frac{G_{en}}{d_3}\right),$$

where $t_{p\ell}$--duration of trip for given range ℓ,

$$t_{p\ell} = \frac{1}{V_{cr\ell}}.$$

In (5.30) the expression $\sum\limits_{\ell} C_{sth\,\ell} N_{so\,\ell}$ comprises the hourly expenditures on maintenance totaled for all distances. After multiplying by the average annual accrued flying time T_g, we obtain the annual expenditures on aircraft maintenance. The method for calculating the amount of current expenditures on engine maintenance C_{dtt} is analogous to the aircraft calculation method.

The wages of flight personnel during year t are determined from the formula

$$C_{w\ell ppt} = C_{w\ell pph} T_g N_{sot}, \tag{5.31}$$

where $C_{w\ell pph}$--hourly wages;

$$C_{w\ell pph} = \sum_{i=1}^{n} K_i a G_o,$$

where K_i--wage coefficient for crew members in relation to the wages of captain taken as one;

n--number of crew members.

Fuel expenditures in year t P_{ft} are determined by the formula

$$P_{ft} = T_g \frac{N_{sot}}{N_{s\Sigma}} P_{kgf} \sum_{\ell} \frac{G_{fcon\,x\,\ell p}}{\ell} V_{cr\ell} N_{ct,\ \ell}, \tag{5.32}$$

where P_{kgf}--price of 1 kg of fuel;

$G_{fconx}\ \ell$--fuel consumption in flight over distance ℓ.

Other expenditures in operation during year t (expenditures related to the training of flight personnel) are determined by the set coefficient for other expenditures K_{oth} according to the formula

$$C_{otht} = [(P_{s,\ drt} + C_{s,\ dtt} + C_{w\ell ppt}] K_{oth}. \tag{5.33}$$

Airport expenses in year t of operation are set by the formula

$$C_{apt} = C_{aph} N_{sot} T_g, \tag{5.34}$$

where C_{aph}--hourly airport expenses:

$$C_{aph} = (c + dG_{com})/t_{pt};$$

here c and d--equation parameters;

t_{pt}--average duration of trip in year t with length L.

In determining the economic effectiveness of the BTS, including the air transport systems, for the criterion (3.24) the least studied question is that of determining the amounts of capital investments into prototype and serial production as well as into operations. Let us give certain recommendations on solving the given questions.

Capital investments into prototype during year t may be calculated in the following manner. Initially one determines the capital investments needed by prototype production for developing the new article $K_{jniokr\Sigma}$. This amount can be determined by a direct calculation. However, in the stage of system forecasting, it is essential also to forecast the amount of capital investments for NIOKR. In the elaborated model for the ATS effectiveness criterion, it has been proposed that the amount $K_{jniokr\Sigma}$ be determined in fractions of the cost $C_{jniokr\Sigma}$:

$$K_{jniokr\Sigma} = C_{jniokr\Sigma} K_{jkniokr}, \qquad (5.35)$$

where $K_{jkniokr\Sigma}$--coefficient expressing ratio of amount of capital investments into development of system to the cost of the NIOKR for the system.

The amount of the coefficient $K_{jkniokr\Sigma}$ is set on the basis of processing the statistical data and this is the cause of the shortcomings in the proposed calculation method. There is no doubt that here modern forecasting methods must be employed for forecasting the amount of the capital investment coefficient. At present the various industrial sectors are at work on forecasting the NIOKR expenditures. However, as a rule, all this work is devoted to just the current NIOKR expenditures $C_{niokr\Sigma}$. At the same time, the amount of capital investments into the NIOKR in a majority of instances, particularly for fundamentally new types of technology, is in no way less than the amount of current expenditures. Obviously, we must be concerned with working out methods for forecasting capital investments into the NIOKR. At first it would be possible to obtain a dependence for the amount of capital investments as a function of the system's parameters X of the following type:

$$K_{jniokr\Sigma} = b \prod_{i=1}^{n} X_i^{b_i}.$$

Capital investments into NIOKR in year t $K_{jkniokrt}$ may be determined from their total amount $K_{jniokr\Sigma}$ using a coefficient of capital investments dynamics $\eta_{jkniokrt}$ indicating what percent of the total amount of capital investments is made in year t.

The amount of capital investments into the serial production of element j of the system in year t K^{n}_{jmt} can be determined by a method analogous to the one used in prototype production, that is, initially one determines the total amount of capital investments and then their annual amount using the coefficients for capital investment dynamics in serial production. The total amount of capital investments is determined from the rates of proportional capital investments.

The capital in manufacturing the system's element j in year t can be determined as the product of multiplying the annual program for the system's element j (aircraft, engines and so forth) by the coefficient for the capital intensiveness of manufacturing K_{jcm}:

$$F_{jmt} = N_{jmt} K_{jcm}. \qquad (5.36)$$

Capital investments into operations are expenditures on new aircraft and engines (not considering expenditures going to replace worn out aircraft and engines) and capital investments into airport facilities K_{it}. The latter are determined in the following manner:

$$K_{it} = N_{s\Sigma}(\overline{P}_s + n_d\overline{P}_d)K_{ci}n_{cit},$$

where K_{ci}--proportional capital investments into airport facilities (in the given case it is considered that they have been calculated in relation to the cost of the aircraft and engine fleet);

n_{cit}--coefficient expressing capital investments dynamics into airport facilities.

5.3. Approximate Method for Determining National Economic Effect (From the Example of the ATS)

The approximate method differs from the above-examined dynamic method in the composition of the expenditures considered. Since at present there are no methods for determining the amount of capital investments into prototype design work K_{niokr} in production related to the creation of new technology K_m and K_{sp}, these expenditures cannot be accounted for with the approximate calculation method. Naturally as these methods and the proportional capital investment rates are worked out these expenditures must be considered in the approximate method.

In the approximate method the system variations are compared not from the condition of an equal volume of work over the forecasted or planned period, as in the case of the dynamic method, but from the condition of the same volume of work A_g which is carried out by the system during the year of its developed operation. The annual volume of transport work for the ATS is determined on the basis of the following data:

1) The length of the routes $L = \{\ell\}$ on which the given system will be operated;

2) The volume of air shipments in reduced cargo tons G_ℓ calculated for a year of developed operation of the system for each distance $\ell \in L$.

The ATS variations are compared with a fixed amount of the system's specific effect. The specific effect is the volume of air transport work A_g in tons of cargo and is calculated in the following manner:

$$A_g = \sum_\ell G_\ell. \tag{5.37}$$

On the basis of the given annual volume of air shipments and the range of parameters characterizing the quality of the designated system variations (for the ATS this is G_{com}, $L_{ca\ell}$), the required aircraft fleet can be determined using the formula:

$$N_{s\Sigma} = \sum_\ell \frac{G_\ell}{K_z G_{com\ell} V_{cr\ell} T_g}, \tag{5.38}$$

where K_z--aircraft load factor.

For the calculated number of aircraft one determines the required number of engines using the formula

$$N_{d\Sigma} = N_{s\Sigma} n_d \frac{T_{s\Sigma}}{\tau_{d\Sigma}} . \tag{5.39}$$

The presence of technical parameters for the system elements and quantitative data on the demand for them make it possible for the methods given in Chapter 4 to determine the cost parameters of the system, as follows: expenditures on NIOKR, the cost of manufacturing and operating the system which, in turn, as component elements becomes part of the ATS economic effectiveness criterion.

The system's economic effectiveness criterion with the approximate calculation method is the minimum of reduced expenditures for the annual volume of its work:

$$Z = \frac{1}{D_c}\Bigg\{\frac{1}{\sum\limits_{t=1}^{T_c}(1+E)^{-t}}\sum_{t=1}^{T_c}\left(C_{cot}+\sum_{j=1}^{n} n_j C_{jot}\right)(1+E)^{-t} + E_n\Bigg[\frac{Z_{cokr}}{N}(1+E)^{T_c Z_{okr-o}} + \sum_{j=1}^{n} n_j \frac{Z_{jokr}}{N_j}(1+E)^{T_j Z_{okr-o}} + P_c + n_j \sum_{j=1}^{n} P_j + K_l)\Bigg\} A_g \rightarrow \min, \tag{5.40}$$

where D_c--annual productivity of system's central element, ton-km per year;
T_c--service life of system central element;
c--system central element;
j--element of system central element;
m_j--number of considered elements of system central element;
C_{cot}--expenditures on operating system central element during year t;
C_{jot}--expenditures on operating system element in year t;
Z_{cokr}--expenditures on working out system central element, rubles;
N_c--the size of manufacturing batch of central element, pieces;
$T_c Z_{okr-o}$--lead in making expenditures on OKR for central element to the system's operation, years;
Z_{jokr}--expenditures on working out system element;
M_j--size of manufacturing batch of system element;
$T_j Z_{okr-o}$--lead in making expenditures on OKR in relation to system's operation, years;
P_c--price of system central element, rubles;
P_j--price of system element;
K_l--other capital investments into operation of system (not counting cost of system elements).

The expression (5.40) can also be given in the following form:

$$Z = Z_{yd}A_g, \qquad (5.41)$$

where Z_{yd}--reduced expenditures per unit of system work;

$$Z_{yd} = C_{yd} + E_n K_{yd}, \qquad (5.42)$$

where C_{yd}--proportional cost of a unit of work;
K_{yd}--capital investments per unit of work.

The calculations using the formula of (5.40) provide a differentiation of expenditures for the individual aircraft elements: engines, equipment and so forth. In existing practices calculations are isolated for just one element, the engines, and for this reason we will divide the aircraft conditionally into two elements: a) engines; b) all remaining aircraft elements.

Let us designate the aircraft by the index ℓa, the first element, the engines, by the index d and the remaining elements by $p\ell$. Considering this (5.40) will assume the form

$$Z = \frac{1}{D_{la}}\left\{\frac{1}{\sum\limits_{t=1}^{T_{la}}(1+E)^{-t}}\sum_{t=1}^{T_{la}}(C_{plot} + n_d C_{dot})(1+E)^{-t} + E_n\left[Z_{plokr}(1+E)^{T_{plZ_{okr-o}}} + n_d Z_{dokr}(1+E)^{T_{dZ_{okr-o}}} + P_{pl} + n_d P_d (1+k_o) + K_i\right]\right\}A_g = \min, \qquad (5.43)$$

where $D_{\ell a}$--annual aircraft productivity;
$T_{\ell a}$--aircraft life;
$C_{p\ell ot}$--expenditures on operation of aircraft (without engines) in year t;
C_{dot}--expenditures on operation of engine in year t;
n_d--number of engines per aircraft;
$Z_{p\ell okr}$, Z_{dokr}--expenditures on the development of the aircraft and engine, respectively, for the manufactured aircraft (without the engine) and the engine;
$T_{p\ell Z_{okr-o}}$, $C_{dZ_{okr-o}}$--lag in making expenditures on OKR in relation to operation;
$P_{p\ell}$, P_d--price of aircraft (without engines) and of engine;
k_o--coefficient for replacement stock of engines in operation;
K_i--capital investments into fixed facilities in operation per aircraft, rubles/units.

In comparing several system variations the expenditures are totaled either for the longest service life $T_{\ell a}$ of all the examined variations or for the adopted calculation period. If it is accepted that the current annual expenditures in operation are constant over the service life of the aircraft, that is, $C_{p\ell ot} = C_{p\ell o} = \text{const.}$

and $C_{dot} = C_{do} = \text{const.}$ [where C_{plo} and C_{do}--annual expenditures on the operation of the aircraft (without the engine) and of the engine], then under this condition the criterion of national economic effectiveness of the ATS will assume the form

$$Z = \frac{1}{D_{la}} \left\{ C_{plo} + n_d D_{do} + E_n \left[Z_{plokr} (1+E)^{T_{pl} Z_{okr-o}} + n_d Z_{dokr} (1+E)^{T_d Z_{okr-o}} + P_{pl} + n_d P_d (1 + k_o) + K_i \right] \right\} A_g = \min. \qquad (5.44)$$

The formula (5.44) considers the full total of one-shot expenditures and the costs of the entire annual volume of the system's work. Along with such indicators it is also possible to use proportional indicators, that is, the proportional one-shot expenditures and the proportional cost of a unit of aircraft work. In this instance the national economic effectiveness criterion is the minimum of reduced expenditures per ton-km:

$$Z_{tkm} = C_{tkm} + E_n K_{tkm},$$

where C_{tkm}--the cost of 1 ton-km;

K_{tkm}--capital investments per ton-km.

From formula (5.43) it can be seen that

$$C_{tkm} = \frac{1}{D_{la} \sum_{t=1}^{T_{la}} (1+E)^{-t}} \sum_{t=1}^{T_{la}} (C_{plot} + n_d C_{dot})(1+E)^{-t}; \qquad (5.45)$$

$$K_{tkm} = Z_{plokr} (1+E)^{T_{la} Z_{okr-o}} + n_d Z_{dokr} (1+E)^{T_d Z_{okr-o}} + P_{pl} + n_d P_d (1 + k_o) + K_i. \qquad (5.46)$$

If it is assumed that current expenditures in operation are fixed, then

$$C_{tkm} = \frac{C_{plo} + n_d C_{do}}{D_{la}}; \qquad (5.47)$$

$$Z_{tkm} = C_{tkm} + \frac{E_n}{G_{com} V_{cr} T_g k_z} \left[Z_{plokr} (1+E)^{T_{la} Z_{okr-o}} + n_d Z_{dokr} (1+E)^{T_d Z_{okr-o}} + P_{pl} + n_d P_d (1 + k_o) + K_i \right] = \min. \qquad (5.48)$$

where k_z--aircraft load factor.

The method for calculating the amount C_{tkm} is generally known and for this reason we would point up the features in calculating the other components of formula (5.48). The capital and current expenditures on NIOKR, the leads between the various stages in the life cycle of the system elements and the amount of capital investments into operations are determined from norms which, in turn, are the result of corresponding forecasts or long-range plan calculations.

The price of a system element (airframe, engine) is determined by the formula

$$P_{jm} = \bar{C}_{jm} + R_n\bar{F}_{jm},$$

where C_{jm}--the average aggregate cost from batch N_j for the system element;
R_n--normed profitability;
$\bar{F}_{jm}$--average amount of capital intensiveness for system element.

In the absence of norms for the capital intensiveness of system elements, their price is determined by the formula

$$P_{jm} = \bar{C}_{jm}(1 + R_c),$$

where R_c--profitability rate calculated for cost of element j.

In determining the national economic effect from the operation of a BTS, a calculation is made of the annual and full effects and the effectiveness level.

The annual national economic effect from the operation of a BTS is calculated from the formula

$$Y_g = (Z_{byd} - Z_{nyd})A_g, \tag{5.49}$$

where Z_{byd} and Z_{nyd}--proportional expenditures for base and new variations.

The full effect over the system's operation period is:

$$Y_\Sigma = E_g \sum_{t=1}^{T_o} (1 + E)^{-t}. \tag{5.50}$$

The level of national economic effectiveness for the new system is:

$$E_r = \frac{C_{byd} - C_{nyd}}{K_{nyd} - K_{byd}}. \tag{5.51}$$

In terms of ATS, the annual national economic effect from the operation of the ATS is calculated from the formula

$$Y_g = (Z_{btkm} - Z_{ntkm})A_g, \tag{5.52}$$

where Z_{btkm}, Z_{ntkm}--expenditures per ton-km for the base and new variations.

The annual effect from the operation of one aircraft is calculated in the following manner:

$$Y_{g\ell a} = (Z_{btkm} - Z_{ntkm})A_{g\ell a}, \tag{5.53}$$

where $A_{g\ell a}$--annual volume of work of aircraft, tkm.

The full effect of a new aircraft is

$$Y_{\ell a} = Y_{g\ell a} \overset{T_{\ell a}}{\underset{t=1}{}} (1+E)^{-t}. \tag{5.54}$$

The level of national economic effectiveness for the new ATS is determined by the formula:

$$E_r = \frac{C_{btkm} - C_{ntkm}}{K_{ntkm} - K_{btkm}}, \tag{5.55}$$

where K_{ntkm}, K_{btkm}--proportional (per ton-km) capital investments for the new and base variations, rubles per year per tkm.

The proportional capital investments for the new variation of the system are:

$$K_{ntkm} = \frac{1}{G_{com}V_{cr}T_g k_z}\left[Z_{p\ell okr}(1+E)^{T_{\ell a}Z_{okr-o}} + n_d Z_{dokr}(1+E)^{T_d Z_{okr-o}} + P_{p\ell} + n_d P_d(1+k_o) + K_i\right] = \min.$$

If as the base variation one uses the system put into operation, the amount of the proportional capital investments put into the base system is determined in the following manner

$$K_{btkm} = \frac{1}{G_{com}V_{cr}T_g k_z}\left[P_{p\ell} + n_d P_d(1+k_o) + K_i\right]. \tag{5.56}$$

5.4. Dynamic Method for Determining Cost Accounting Effect of BTS

The cost accounting effect of a new BTS is the savings in expenditures from its production and use in comparison with the BTS adopted as the basis for comparison with the set volume and quality of product. The cost accounting effect is apparent for the manufacturer and user of the system, and for this reason it is possible to differentiate the annual effects for the system's consumer and manufacturer and the full effect over the period of producing and operating the new system.

The annual cost accounting effect for the manufacturer from producing the new BTS is determined on the basis of the criterion of (3.26) from the formula

$$Y_{cmt} = \sum_{j=1}^{m} \{P_{njmt} - (C_{njmt} + E_{nm}K_{njmt})] - [P_{bjmt} - C_{bjmt} + E_{nm}K_{bjmt})]\} \qquad (5.57)$$

or

$$Y_{cmt} = \sum_{j=1}^{m} [(P_{njmt} - C_{njmt}) - (P_{bjmt} - C_{bjmt}) - E_{nm}(K_{njmt} - K_{bjmt})]. \qquad (5.58)$$

If we designate the amount $(K_{njmt} - K_{bjmt})$ by ΔK_{jmt}, then the formula assumes the form

$$Y_{emt} = \sum_{j=1}^{m} [P_{njmt} - C_{njmt}) - (P_{bjmt} - C_{bjmt}) - E_{nm}\Delta K_{jmt}], \qquad (5.59)$$

where P_{njmt}--the price of manufacturing element j of the new system in year t;
C_{njmt}--the cost of manufacturing element j of the new system in year t;
P_{bjmt}, C_{bjmt}--price and cost of manufacturing element j of base system in the year t;
E_{nm}--the normed coefficient of capital investment effectiveness in the manufacturing sector;
ΔK_{jmt}--additional capital investments for manufacturer in year t (increase in amount of capital for manufacturer in year t).

In an analogous manner we determine the cost accounting effect for the consumer (in operating the system:

$$Y_{eot} = (P_{not} - C_{not}) - (P_{bot} - C_{bot}) - E_{no}\Delta K_{ot}, \qquad (5.60)$$

where P_{not}, C_{not}--price and cost of operating new system in year t;
P_{bot}, C_{bot}--price and cost of operating base system in year t;
E_{no}--normed coefficient of capital investment effectiveness in the consuming sector;
ΔK_{ot}--additional capital investments in operations during year t (increase in amount of capital for consumer in year t).

The full cost accounting effect in the year t will be:

$$Y_{et} = Y_{emt} + Y_{eot}. \qquad (5.61)$$

The full cost accounting effect over the period of producing and operating the system is determined by totaling the annual effects:

$$Y_{e\Sigma} = \sum_{t=1}^{T_m} Y_{emt} + \sum_{t=1}^{T_o} Y_{eot}. \qquad (5.62)$$

The basic question in determining the cost accounting effect for the consumer and manufacturer of the system is the setting of prices for the system elements. If in the forecasting stage the price of a system's element in year t is determined from its so-called lower limit from the formula $P_{jmt} = C_{jmt} + E_{nm}F_{jmt}$ (here F_{jmt}--capital intensiveness of the system's element in year t), and the price for the elements of the base system is adjusted in such a manner that it provide the manufacturer with the normed profitability, then the effect for the manufacturer will equal zero and the consumer will receive the entire effect. In actuality, according to formula (5.58) we obtain:

$$Y_{emt} = \sum_{j=1}^{m} [(C_{njmt} + E_{nm} + F_{jmt} - C_{njmt}) - (C_{bjmt} + E_{nm}F_{bjmt} - C_{bjmt}) - E_{nm}(F_{njmt} - F_{bjmt})] = 0. \tag{5.63}$$

In this instance, the full cost accounting effect in the year t will be $Y_{et} = Y_{eot}$.

If it is accepted that the prices in operating the system remain fixed, that is, $P_{not} = P_{dot} = \text{const.}$, then the full cost accounting effect of the system in year t can be determined by the formula

$$Y_{et} = (C_{bot} - C_{not}) - E_{no}\Delta K_{ot}. \tag{5.64}$$

The full cost accounting effect of the system over the period of operation is:

$$Y_{e\Sigma} = \sum_{t=1}^{T_o} [(C_{bot} - C_{not}) - E_{no}\Delta K_{ot}]. \tag{5.65}$$

The level of cost accounting effectiveness for the new BTS is determined by the indicator for the effectiveness level of additional capital investments

$$E_r = \frac{\sum_{t=1}^{T_o} [(P_{not} - C_{not})(P_{bot} - C_{bot})]}{\sum_{t=1}^{T_o} \Delta K_{ot}}. \tag{5.66}$$

The amount of the indicator E_r is compared with the sectorial normed coefficient of capital investment effectiveness.

The indicator E_r indicates the level of the new system's effectiveness in comparison with the old system. Along with the effectiveness level also of interest is the indicator of the general effectiveness of the new system, the effectiveness level of the full capital investments:

$$Y_{cf} = \frac{\frac{1}{T}\sum_{t=1}^{T_o}(P_{not} - C_{not})}{K_{no}}, \qquad (5.67)$$

where K_{no}--full amount of capital investments into operations under new system.

The amount of the indicator Y_{cf} is compared with the sectorial standard and the analogous indicator for the base system.

5.5. Approximate Method for Determining Cost Accounting Effect (From the Example of an ATS)

The criterion for the cost accounting effectiveness of an ATS can be obtained by the full and proportional expenditures.

In the first instance, in accord with the formula (3.26), the criterion will be

$$Y_e = (P_{no} - C_{no}) - (P_{bo} - C_{bo}) - E_{no}\Delta K_o = \text{max.}, \qquad (5.68)$$

where P_{no}, C_{no}, P_{bo}, C_{bo}--price and cost of average annual volume of work for new and base systems;
ΔK_o--additional capital investments into operation.

If one uses proportional indicators, then the criterion will be

$$Y_{etkm} = (P_{ntkm} - C_{netkm}) - (P_{dtkm} - C_{detkm}) - E_{no}\Delta K_{etkm} = \text{max.}, \qquad (5.69)$$

where P_{ntkm}, P_{dtkm}--price of 1 ton-km for new and base systems;
C_{netkm}, C_{betkm}--the cost of 1 ton-km for new and base systems;
ΔK_{etkm}--additional annual capital investments per ton-km.

For the ATS, the most typical case is $P_{no} = P_{bo}$. In this instance the equations of (5.68) and (5.69) can be rewritten:

$$Y_e = (C_{bo} + E_{on}K_{bo}) - (C_{no} + E_{on}K_{no}) = \text{max.}, \qquad (5.70)$$

$$Y_{etkm} = (C_{betkm} + E_{on}K_{betkm}) - (C_{netkm} + E_{on}K_{netkm}) = \text{max.} \qquad (5.71)$$

The expressions in parentheses are the reduced cost accounting expenditures which can be designated by Z_{betkm} and Z_{netkm}.

The annual cost accounting effect of a new ATS is calculated from the formula (5.68). If we use proportional expenditures and also assume that the prices for transport services are fixed, then the annual cost accounting effect in operations is:

$$Y_{eg} = (Z_{betkm} - Z_{netkm})A_g, \qquad (5.72)$$

where Z_{betkm}, Z_{netkm}--cost accounting reduced expenditures per ton-km for base and new variations.

The method of calculating the cost of a ton-km is generally known. The amount of the proportional capital investments is determined from the formula

$$K_{etkm} = \frac{1}{G_{com}V_{cr}T_{g}k_{z}} [P_{\ell a} + n_{d}P_{d}(1 + k_{o}) + K_{i}]. \tag{5.73}$$

The full cost accounting effect over the period of operating the systems is calculated by the formula

$$Y_{e\Sigma} = Y_{eg}T_{o}. \tag{5.74}$$

The annual cost accounting effect from the operation of one aircraft is:

$$Y_{eg\ell a} = (Z_{betkm} - Z_{netkm})A_{g} \quad . \tag{5.75}$$

The full effect of the new aircraft is:

$$Y_{\Sigma\ell a} = Y_{g\ell a}T_{\ell a}, \tag{5.76}$$

where $T_{\ell a}$--service life of aircraft in years.

The effectiveness level of additional capital investments is:

$$E_{r} = \frac{C_{betkm} - C_{netkm}}{K_{netkm} - K_{betkm}} \ . \tag{5.77}$$

The effectiveness level of full capital investments is:

$$Y_{cf} = \frac{P_{tkm} - C_{etkm}}{K_{etkm}} \ , \tag{5.78}$$

where P_{tkm}--price of one ton-km.

5.6. Determining the Limit Price of an Aircraft in the Stage of Its Development

In the feasibility studies for a new aircraft, along with the calculations for the economic effect and the economic effectiveness level, its limit price is also set. The limit price is a compulsory parameter of a technical requirement for the designing of an aircraft and is determined in the following manner.

The upper price limit for an aircraft is derived from the conditions of equal advantage from the operation of the new and base aircraft, that is, the annual effect in operations is reduced to zero. With the given condition, according to formula (5.72)

$$Z_{netkm} = Z_{betkm}$$

or

$$C_{netkm} + E_{no}K_{etkm} = C_{betkm} + E_{no}K_{betkm}. \tag{5.79}$$

Having brought out the content of the individual components of equation (5.79), we obtain

$$\frac{1}{X_{nh}}\left(\frac{P_{\ell avp}}{\tau_{n\Sigma\ell a}}+\frac{P_{\ell avp}\gamma_{r\ell a}}{\tau_{n\ell a}}+I_{nh}+E_{no}\frac{P_{\ell avp}}{T_g}\right)=$$

$$\frac{1}{X_{bh}}\left(\frac{P_{b\ell a}}{\tau_{b\Sigma\ell a}}+\frac{P_{b\ell a}\gamma_{r\ell a}}{\tau_{b\ell a}}+I_{bh}+E_{no}\frac{P_{b\ell a}}{T_g}\right),\qquad(5.80)$$

where X_{nh}, X_{bh}--productivity per hour for new and base aircraft, tkm/hour;
$\gamma_{r\ell a}$--expenditures on one aircraft overhaul in fraction of its price;
I_{nh}, I_{bh}--expenditures on one hour of operation of the new and base aircraft (without renovation deductions), rubles/hour.

In solving equation (5.80) for $P_{\ell avp}$, we obtain

$$P_{\ell avp}=\frac{\frac{X_{nh}}{X_{bh}}\left(\frac{P_{b\ell a}}{\tau_{b\Sigma\ell a}}+\frac{P_{b\ell a}\gamma_{\ell ar}}{\tau_{b\ell a}}+I_{bh}+E_{no}\frac{P_{b\ell a}}{T_g}\right)-I_{nh}}{\frac{1}{\tau_{n\Sigma\ell a}}+\frac{\gamma_{r\ell a}}{\tau_{n\ell a}}+\frac{E_{no}}{T_g}}.\qquad(5.81)$$

The limit price for the aircraft is calculated by the formula

$$P_{\ell a\ell}=P_{\ell avp}B,\qquad(5.82)$$

where B--reduction factor related to a decline in production outlays for aircraft as a result of developing their serial output.

5.7. Estimating Economic Effectiveness of BTS Functional Elements

Above we have examined the criteria and methods for an economic estimate of the BTS. The BTS have a complex hierarchical structure and include a large number of elements. The question arises of the criteria and methods for estimating the individual elements and their connection to the criteria and methods for estimating the system as a whole.

In the literature there are two approaches to the question of an economic estimate of system functional elements: 1) a functional element is viewed as an independent object and in its economic estimate there is a restriction to basic consideration of the technical and cost parameters for it alone; 2) the functional element is viewed in an organic relationship to the system and the element's effectiveness estimate is based on a consideration of the system's functional and technical-economic indicators which can be obtained in using elements with different parameters in the system.

Let us examine the first approach to an economic estimate of a functional element from the example of an aviation engine. Modern aviation gas turbine engines (GTD) are characterized by a broad range of possible parameter values such as thrust, air consumption, specific fuel consumption, specific engine weight and so forth. The entire multiplicity of engine parameters can be reduced to the following groups: specific, limiting, technical and economic.

An example of the engine's specific property is its capacity to develop thrust which is used by the aircraft for carrying out its specific property, the ability to transport cargo.

Among the limiting parameters of an engine are: heat resistance, frost resistance, noise level, the degree of polluting the environment with fuel combustion products, maximum speed, maximum altitude, fire safety and so forth.

In addition to the specific and limiting properties, an engine has properties which in their aggregate determine the consumption of resources related to its production, storage, transporting and operation. These properties are determined by the technical parameters of an engine (the degree of increasing pressure in the compressor, gas temperature ahead of the turbine, the degree of pressure rise in the fan and so forth) and by economic parameters, for example, by price, by manufacturing labor intensiveness, by engine material intensiveness and so forth. The engine's technical and economic parameters determine its technical level and the level of its quality and are divided into natural and cost. An example of the natural indicators for engine quality level is: the specific weight γ (the ratio of engine weight to its thrust), specific fuel consumption C_R (the ratio of fuel consumption per hour to the amount of thrust), the specific frontal area f_{fro} (the ratio of the area of the engine's maximum cross-section to the amount of thrust). If the weight, overall dimensions and fuel consumption are viewed as expenditures, then it turns out that the reduced proportional or specific parameters characterize the individual aspects of engine efficiency. The less weight, volume or fuel which must be spent to achieve a certain thrust with a certain dependability level, the more efficient an engine is in operating it on an aircraft. The range, altitude, speed of flight and other indicators of aircraft quality depend upon the amount of these indicators.

In assessing an engine, cost indicators can also be used, for example, the indicators for the proportional expenditures of serial production: C_R--the proportional price of 1 kg of engine thrust; $P_{m,\ R}$--proportional material expenditures per kg of engine thrust; $C_{w,\ R}$--proportional wage expenditures of production workers per kg of engine thrust. The list of the proportional cost indicators for engines could be broadened. The aggregate of such proportional indicators sufficiently describes an engine as an object of production.

Of all the listed indicators, the most widespread is the indicator for the cost of 1 kg of thrust which is calculated by the formula

$$P_R = P_d:R,$$

where P_d--engine price;
R--engine thrust.

Let us attempt to transform this indicator by incorporating other engine parameters into it. If we introduce the operating life indicator, then we obtain an indicator for engine cost per kg of thrust in a unit of time of work (hour) $P_d:(R\tau_{d\Sigma})$.

It is also possible to broaden the scope of considered expenditures and along with engine price consider operating expenditures: expenditures on major overhaul, expenditures on fuel and maintenance and so forth. Then we will obtain an indicator which expresses the proportional cost of 1 kg of engine thrust per hour:

$$C_{dhR} = \frac{P_d + \sum_{1}^{n} P_{dr} + C_{dm\Sigma} + P_{f\Sigma}}{R\tau_{d\Sigma}},$$

where P_{dr}--the price of one major overhaul of an engine;
$C_{dm\Sigma}$--expenditures on engine maintenance over period $\tau_{d\Sigma}$;
$P_{f\Sigma}$--price of fuel consumed by engine in $\tau_{d\Sigma}$.

The indicator C_{dhR} can be given in the form:

$$C_{dhR} = C_{dh}:R,$$

where C_{dh}--cost of 1 hour of engine operation.

The indicator C_{dhR} considers a number of engine properties: thrust, service life, specific weight, specific fuel consumption and specific frontal area. However, certain engine properties are not fully considered by this indicator. The specific weight and specific frontal area are reflected in the cost and price of the engine while fuel consumption is shown in operating expenditures. However these engine properties are not sufficiently considered in just the engine price and its operating expenditures. The specific weight of an engine, the specific frontal thrust and fuel consumption influence the load factor of the aircraft and its transport effectiveness. If the engine is viewed separately from the aircraft, then in an economic estimate these engine properties cannot be accounted for.

In an economic estimate of an engine, for any stationary unit these properties can be disregarded. But for aviation engines, the weight, frontal area and fuel consumption are important properties which must be considered in their economic estimate and for this reason for a full economic estimate of an aviation engine it must be viewed in relation to the aircraft.

Thus, in an economic estimate of an aircraft engine, as an independent object, it is possible to use a number of proportional cost indicators of which the most general is the price of 1 kg of thrust developed by the engine per hour. However, with such an approach the engine estimate is incomplete, as it cannot consider a number of engine properties (γ, C_R, f_{fro}) which are manifested in relation to the aircraft. For this reason the fullest estimate of engine economic effectiveness can be achieved in the aircraft--engine system.

Let us recall that an aircraft engine here is viewed as an example of a system functional element. For this reason it can be concluded that a full estimate of a system functional element can be given if the functional element is viewed in an inseparable link with the system. However, let us return to the example of the aircraft engine.

One other attempt can be made to fully estimate such engine properties as specific frontal area and specific fuel consumption. For this, in the indicator C_{dhR} it is possible to consider not the entire amount of engine thrust but only its useful or effective part.

Effective thrust R_{eff} is determined from the full engine thrust minus the thrust used to overcome the drag of the engine's R_{cy} and the moving of the intrinsic weight R_G as well as fuel R_{Gf}:

$$R_{eff} = R - R_{cy} - R_G - R_{Gf}.$$

Engine drag is determined by the formula

$$R_{cy} = C_{xcy} \frac{\rho v^2}{2} F_{mid},$$

where C_{xcy}--drag factor of engine nacelle;
F_{mid}--the engine cross-section area, m^2;
ρ--air density.

The thrust consumed on moving the engine and fuel in flight is calculated by the formula:

$$R_G = G \frac{1}{K_n}; \quad R_{Gf} = G_f \frac{1}{K_n},$$

where G--engine weight;
G_f--weight of fuel on aircraft;
K_n--aerodynamic quality of airframe;

$$K_n = C_y:C_{xp},$$

where C_y--wing lift coefficient;
C_{xp}--drag coefficient of airframe (without power unit) under cruising flight conditions.

Thus, if we incorporate in the indicator C_{dhR} the effective thrust R_{eff} in the place of thrust R, then the indicator of the engine's economic estimate will be the amount $C_{dhR_{eff}}$. The indicator $C_{dhR_{eff}}$, in comparison with C_{dhR}, better reflects certain engine properties (γ, C_R, f_{fro}). However, this happens precisely because the given indicator considers a number of important aircraft indicators along with the engine indicators. Such an indicator can be used only in approximate calculations, as it does not make it possible to fully consider the influence of engine properties on the economy of the aircraft. For a full economic estimate of an engine, it must be viewed in a single aircraft--engine system in which engine parameters depend upon the aircraft class and the conditions of its operation.

As a rule, engines are created for a specific aircraft and their economic estimate is made on the basis of the given aircraft. If it is necessary to compare engines independently of a specific aircraft (sometimes aircraft shortcomings can have a negative impact on the estimate of a good engine), then in this instance one may use hypothetical aircraft which can be created on the basis of the compared engines with a certain development level of science and technology. It must be emphasized that the engines are estimated while the aircraft are a means of estimation. For comparing several engine variations, the expenditures related just to the aircraft remain fixed and do not influence the comparison results. In the comparison consideration is given only to those expenditures and effects which are related to the quality features of the compared engines.

Engines can be also viewed as a functional element of an aircraft system. Naturally just as there is a relationship between an element and a system there is a relationship between the criteria for estimating them. The main thing related to the criteria for the element and the system is that the element estimate criterion should reflect the degree to which the adopted partial solutions (the solutions relating to the element) correspond to the interests of the system's general goals. It is very difficult to solve such a problem in elaborating the particular criteria for an element.

For functional elements, this can be solved in the following manner. If one views the system's hierarchy, then the criterion for estimating the system itself can be the criterion for estimating any element. Strictly speaking, the system's criterion is not the criterion for estimating an element. Here, in essence, it is a question of estimating several variations of a system which differ in the parameters of the elements comprising them. In assessing the various variations of a system, we thereby give an estimate to the different variations of the elements. Experience has shown that this approach to estimating aircraft engines has proven effective. Possibly, to a significant degree this is explained by the fact that in the effect and in the expenditures for an aircraft system the engines hold a large proportional amount. Suffice it to say that in the price of an aircraft, the share of engine costs approaches 25 percent, and in aircraft operating expenditures, the proportional amount of expenditures related to the creation and operation of engines approaches 50 percent.

Thus, if we adopt the above-given ideas on estimating system elements to the engines, the full estimate of the engines can be made only within the aircraft--engine system using the criterion for estimating this system. The aircraft--engine system, in turn, can be viewed as a subsystem of the higher level systems: in the aggregate of a certain class and type of aircraft; in the air transport system; in the unified national transport system, and so forth up to the national economy as a whole.

For decision taking on each level of examination, it is essential to have a corresponding hierarchical system of criteria and models describing the structure and dynamics of the given level phenomenon. If one examines such an element of a system as aviation engines, it is possible to establish two basic classes of problems related to its economic estimate:

1) Determining the optimum combinations of different types of aircraft and engines, that is, the problem of optimizing the structure and composition (quantity) of the aircraft and corresponding engine fleet; this problem is examined on the level of the aircraft engine fleet for civil aviation;

2) Determining the optimum engine parameters in terms of the economic criterion with the set basic aircraft characteristics; such problems correspond to the level of the aircraft--engine system.

In the first class are the problems of long-range planning for technical progress in aircraft engine construction, that is: determining the optimum parametric series of aircraft and engines; determining the economic effectiveness from the introduction of a new type of engine or modification; determining the additional effect or loss from reducing (increasing) the development time and the earlier (later) beginning for the development and operation of a new type of engine or modification; modeling the forecasting problems for economically sound moments to replace generations of aviation equipment.

In the second type are the problems of determining economic effectiveness for variations of the same type and purpose of engine and the optimization of its parameters for economic criteria with the set basic aircraft parameters.

All the above-listed problems are solved with the help of particular criteria for estimating the economic effectiveness of aircraft engines: in the system of a certain class and type of aircraft fleet; in the system of the engine fleet.

The most general case is the estimating of an engine within a system of a certain class and type of aircraft fleet. Since the engine is estimated in the aircraft system, here one can fully apply the dynamic and approximate models for estimating the economic effectiveness of the ATS as given in the previous paragraphs of the present chapter.

In order to determine the economic effect from an engine or any other system element, it is essential to allocate the overall effect in the system between the individual elements. In the work [10] it is recommended that the effect be distributed proportionally to the expenditures on wages related to the development and production of the elements and to the coefficients for the scientific and technical significance of the work on element j. Thus, the annual or full economic effect of a new engine (element j of a system) is determined by the formula

$$Y_j = \frac{YZ_{wj}K_{stj}}{\sum_{j=1}^{n} K_{stj}Z_{wj}} \qquad (5.83)$$

where n--number of elements accounted for in system (usually two elements are accounted for: airframe and engines);

Z_{wj}--expenditures on wages related to the development and manufacturing of system element j;

K_{stj}--coefficient for scientific and technical significance of development of element j.

For establishing the design of a system functional element, it is essential to determine its limit price. The limit price of a functional element can be calculated on the basis of the limit price of the system central element. For example, the limit price of a new engine in the development stage is determined on the basis of the limit price of the aircraft.

The upper price limit of an aircraft can be represented in the following form:

$$P_{\ell a vp} = P_{p\ell vp} + n_d P_{dvp}, \tag{5.84}$$

where $P_{\ell a vp}$--upper price limit of airframe;
P_{dvp}--upper price limit for engine.

Let us assume that the upper price limits for the airframe and engine are in the same ratio as their lower limits, that is:

$$P_{p\ell vp} : P_{dvp} = P_{p\ell np} : P_{dnp};$$

$$\gamma_{np} = P_{p\ell np} : P_{dnp}. \tag{5.85}$$

Having substituted the expression for the upper price limit of the aircraft (5.84) in the equation (5.80), we obtain

$$\frac{1}{X_{nh}} \left(\frac{P_{p\ell vp}}{\tau_{n\Sigma p\ell}} + \frac{n_d P_{dvp}}{\tau_{n\Sigma d}} + \frac{P_{p\ell vp}\gamma_{rp\ell}}{\tau_{n\ell a}} + \frac{P_{dvp}\gamma_{rd}}{\tau_{nd}} + I_{nh} + E_{no} \frac{P_{p\ell vp} + n_d P_{dvp}}{T_g} \right) =$$

$$\frac{1}{X_{bh}} \left(\frac{P_{bp\ell}}{\tau_{b\Sigma\ell a}} + \frac{n_d P_{bd}}{\tau_{b\Sigma d}} + \frac{P_{bp\ell}\gamma_{rp\ell}}{\tau_{b\ell a}} + \frac{P_d \gamma_{rd}}{\tau_{bd}} + I_{bh} + E_{no} \frac{P_{bp\ell} + n_d P_{bd}}{T_g} \right). \tag{5.86}$$

In substituting the value $P_{p\ell vp}$ from the equation (5.85) into the equation (5.86) and solving it for P_{dvp}, we obtain

$$P_{dvp} = \frac{\dfrac{X_{nh}}{X_{bh}} \left(\dfrac{P_{bp\ell}}{\tau_{b\Sigma p\ell}} + \dfrac{n_d P_{bd}}{\tau_{b\Sigma d}} + \dfrac{P_{bp\ell}\gamma_{rp\ell}}{\tau_{bp\ell}} + \dfrac{P_{bd}\gamma_{rd}}{\tau_{bd}} + I_{bh} + E_{no} \dfrac{P_{bp\ell} + n_d P_{bd}}{T_g} \right) - I_{nh}}{\dfrac{\gamma_{np}}{\tau_{n\Sigma p\ell}} + \dfrac{n_d}{\tau_{n\Sigma d}} + \dfrac{\gamma_{rp\ell}}{\tau_{np\ell}} + \dfrac{\gamma_{rd}}{\tau_{nd}} + E_{no} \dfrac{\gamma_{np} + n_d}{T_g}}. \tag{5.87}$$

The engine limit price is calculated from a formula analogous to (5.82).

BIBLIOGRAPHY

1. "Materialy XXV S"yezda KPSS [Materials of the 25th CPSU Congress], Moscow, Politizdat, 1976, 256 pp.

2. Akopov, P. L., "Certain Methodological Questions in the Forecasting of Expenditures on OKR and Serial Production of Aircraft Engines," "Tezisy Dokladov na Vsesoyuznoy Konferentsii 'Ekonomicheskiye Problemy Prognozirovaniya Razvitiya Bol'shikh Tekhnicheskikh Sistem'" [Theses of Papers at the All-Union Scientific Conference on "Economic Problems in Forecasting the Development of Large Technical System"], Moscow, Izd. MAI, 1971, pp 32-34.

3. Akopov, P. L., "Calculating the Costs of Increased-Life Machines," STANDARTY I KACHESTVO, No 9, 1968, pp 61-63.

4. Amosov, N. M., "Modelirovaniye Slozhnykh Sistem" [Modeling Complex Systems], Kiev, Naukova Dumka, 1968, 88 pp.

5. Barabash, Yu. L., Varskiy, B. V., Zinov'yev, V. T. et al., "Voprosy Statisticheskoy Teorii Raspoznavaniya" [Questions of Statistical Pattern Recognition Theory], Moscow, Sovetskoye Radio, 1967, 399 pp.

6. Beshelev, S. D., Gurvich F. G., "Matematiko-Statisticheskiye Metody Ekspertnykh Otsenok" [Statistical Mathematics Methods of Expert Estimates], Moscow, Statistika, 1974, 159 pp.

7. Butkovskiy, A. G., "Teoriya Optimal'nogo Upravleniya Sistemami s Raspredelennymi Parametrami" [Optimum Control Theory for Systems With Distributed Parameters], Moscow, Nauka, 1965, 474 pp.

8. Venttsel', Ye. S., Ovcharov, L. A., "Teoriya Veroyatnostey" [Probability Theory], Moscow, Nauka, 1973, 366 pp.

9. Gatovskiy, L. M., "Nauchno-Tekhnicheskiy Progress i Ekonomika Razvitogo Sotsializma" [Scientific and Technical Progress and an Economy of Developed Socialism], Moscow, Nauka, 1974, 431 pp.

10. Golosovskiy, S. I., "Ekonomicheskaya Effektivnost' Nauchnykh Issledovaniy i Razrabotok" [Economic Effectiveness of Scientific Research and Development], Moscow, Moskovskiy Rabochiy, 1975, 167 pp.

11. Grubov, V. I., Ivakhnenko, A. G., Mandrovskiy-Sokolov, B. Yu., "Promyshlennaya Kibernetika. Spravochnik" [Industrial Cybernetics. A Reference Manual], Kiev, Naukova Dumka, 1966, 447 pp.

12. Demidovich, B. P., Maron, I. A., Shuvalova, E. Z., "Chislennyye Metody Analiza" [Numerical Analysis Methods], Moscow, Nauka, 1973, 367 pp.

13. Dobrov, G. M., "Prognozirovaniye Nauki i Tekhniki" [Forecasting Science and Technology], Moscow, Nauka, 1969, 208 pp.

14. "Dolgosrochnyye Programmy Kapital'nykh Vlozheniy. Ekonomicheskiye Problemy i Modeli" [Long-Range Capital Investment Programs. Economic Problems and Models], edited by V. P. Krasovskiy, Moscow, Ekonomika, 1974, 270 pp.

15. Draper, N., Smith, H., "Prikladnoy Regressionnyy Analiz" [Applied Regression Analysis], Moscow, Statistika, 1973, 392 pp.

16. Ezekiel, M., Fox, K. A., "Metody Analiza Korrelyatsii i Regressiy Lineynykh i Krivolineynykh" [Methods of Correlation and Regression Analysis: Linear and Curvilinear], translated from the English, Moscow, Statistika, 1966, 558 pp.

17. Kaspin, V. I., "Factor Analysis in Statistical and Forecast Research," in the book: "Tezisy Dokladov I Konferentsii po Primeneniyu EMM i Sredstv VT v Upravlenii Standartizatsiyey i Kachestvom Produktsii, Metrologiyey i v Spravochno-Informatsionnom Obsluzhivanii" [Theses of Papers from the First Conference on the Use of Mathematical Economics Methods and Computers in Standaridization and Product Quality Control, in Metrology and Information References Services], Moscow, Izd. Gosstandarta, 1971, pp 37-41.

18. Kaspin, V. I., Lisichkin, V. A., "The Concepts of Forecasting," in the collection: "Sorevnovaniye Dvukh Sistem" [The Competition of the Two Systems], Moscow, Nauka, 1971, pp 469-492.

19. Kulikov, N. K., "Multifactor Forecasting on the Basis of Flexible Structure Functions," in the collection: "Osnovnyye Problemy i Zadachi Nauchnogo Prognozirovaniya" [Basic Problems and Tasks of Scientific Forecasting], Moscow, Znaniye, 1972, p 151.

20. Kulikov, N. K., "Elementy Vysshey Matematiki na Osnove Funktsiy s Gibkoy Strukturoy" [Elements of Higher Mathematics on the Basis of Flexible Structure Functions], Moscow, Izd. MTIPP, 1972, 125 pp.

21. Lange, O., Banasinski, A., "Teoriya Statistiki" [Theory of Statistics], translated from the Polish, Moscow, Statistika, 1971, 399 pp.

22. Lisichkin, V. A., "Otraslevoye Nauchno-Tekhnicheskoye Prognozirovaniye" [Sectorial Scientific and Technical Forecasting], Moscow, Ekonomika, 1971, 231 pp.

23. Lisichkin, V. A., "Teoriya i Praktika Prognostiki" [Theory and Practice of Forecasting], Moscow, Nauka, 1972, 224 pp.

24. Lisichkin, V. A., Ratnikova, Ye. I., Kaspin, V. I., "Prognozirovaniye Bibliotechnogo Dela v SSSR" [Forecasting of Library Sciences in the USSR], Moscow, Izd. GB im V. I. Lenin, 1972, 56 pp.

25. Lopukhin, M. M., "PATTERN--Metod Planirovaniya i Prognozirovaniya Nauchnykh Rabot" [PATTERN--A Method of Planning and Forecasting for Scientific Projects], Moscow, Sovetskoye Radio, 1971, 159 pp.

26. Lowly, D. N., Maxwell, A. E., "Faktornyy Analiz kak Statisticheskiy Metod" [Factor Analysis as a Statistical Method], translated from the English, Moscow, Mir, 1967, 144 pp.

27. L'vov, D. S., Rubinshteyn, A. Ya., "Izmereniye Effektivnosti Proizvodstva" [Measuring Production Effectiveness], Moscow, Ekonomika, 1974, 143 pp.

28. Meleshchenko, Yu. S., "Tekhnika i Zakonomernosti Yeye Razvitiya" [Technology and its Developmental Patterns], Leningrad, Lenizdat, 1970, 246 pp.

29. Mesarovich, M., Mako D., Takahara, Y., "Teoriya Iyerarkhicheskikh Mnogourovneykh Sistem" [Theory of Hierarchical, Multi-Level Systems], translated from the English, Moscow, Mir, 1973, 344 pp.

30. Mitropol'skiy, A. K., "Tekhnika Statisticheskikh Vychisleniy" [Statistical Computation Techniques], Moscow, Fizmatgiz, 1961, 480 pp.

31. Rumshinskiy, L. Z., "Matematicheskaya Obrabotka Rezul'tatov Eksperimenta" [Mathematical Processing of Experiment Results], Moscow, Nauka, 1971, 192 pp.

32. Sarkisyan, S. A., Golovanov, L. V., "Prognozirovaniye Razvitiya Bol'shikh Sistem" [The Forecasting of Development for Large Systems], Moscow, Statistika, 1975, 192 pp.

33. Sarkistyan, S. A., Kaspin, V. I., Nechayev, P. A., "Prognozirovaniye Razvitiye Grazhdanskoy Aviatsii" [Forecasting Civil Aviation Development], Minsk, Izd. NIIE pri Gosplane SSSR, 1973, 47 pp.

34. Sarkisyan, S. A., Minayev, E. S., "Ekonomicheskaya Otsenka Letatel'nykh Apparatov" [Economic Evaluation of Aircraft], Moscow, Mashinostroyeniye, 1972, 179 pp.

35. Starik, D. E., Kaspin, V. I., "Prognozirovaniye i Effektivnost' Nauchno-Tekhnicheskogo Progressa" [Forecasting and Effectiveness of Scientific and Technical Progress], Moscow, Izd. MAI, 1975, 61 pp.

36. "Statisticheskoye Izmereniye Kolichestvennykh Kharakteristik" [Statistical Measurement of Quantitative Characteristics], Moscow, Statistika, 1972, 171 pp.

37. Theil, H., "Prikladnoye Ekonomicheskoye Prognozirovaniye" [Applied Economic Forecasting], translated from the English, Moscow, Progress, 1970, 510 pp.

38. Theil, H., "Ekonomicheskiye Prognozy i Prinyatiye Resheniy" [Economic Forecasts and Decision Taking], translated from the English, Moscow, Statistika, 1971, 488 pp.

39. "Teoriya i Praktika Prognozirovaniya Razvitiya Nauki i Tekhniki v Stranakh--Chlenakh SEV" [The Theory and Practice of Forecasting the Development of Science and Technology in the CEMA Countries], Moscow, Ekonomika, 1971, 407 pp.

40. Terekhov, L. L.,"Ekonomika-Matematicheskiye Metody" [Mathematical Economics Methods], Moscow, Statistika, 1972, 360 pp.

41. Timofeyeva, N. M., Skorodumov, S. A., Negru, A. I., "Metodicheskiye Polozheniya Razrabotki Prognozov Razvitiya Tekhnologii Otrasli Mashinostroyeniya" [Procedural Provisions on Working Out Forecasts for the Development of Production Methods in the Machine Building Sector], Minsk, Izd. BelNITI i TE, 1975, 201 pp.

42. Harman, H., "Sovremennyy Faktornyy Analiz" [Modern Factor Analysis], translated from the English, Moscow, Statistika, 1972, 486 pp.

43. Haustein, G., "Metody Prognozirovaniya v Sotsialisticheskoy Ekonomike" [Forecasting Methods in a Socialist Economy], translated from the German, Moscow, Progress, 1971, 398 pp.

44. Himmelblau, D., "Analiz Protsessov Statisticheskimi Metodami" [Process Analysis by Statistical Methods], translated from the English, Moscow, Mir, 1973, 960 pp.

45. Churchman, W., Ackoff, R., Arnow, L., "Vvedeniye v Issledovaniye Operatsiy" [Introduction to Operations Research], translated from the English, Moscow, Nauka, 1968, 486 pp.

46. Ayres, R., "Nauchno-Tekhnicheskoye Prognozirovaniye i Dolgosrochnoye Planirovaniye" [Scientific and Technical Forecasting and Long-Range Planning], translated from the English, Moscow, Mir, 1971, 296 pp.

47. Yule, G. U., Kendal, M. G., "Teoriya Statistiki" [Theory of Statistics], Moscow, Gosstatizdat, 1960, 779 pp.

48. Yampol'skiy, S. M. Lisichkin, V. A., "Prognozirovaniye Nauchno-Tekhnicheskogo Progressa" [Forecasting Scientific and Technical Progress], Moscow, Ekonomika, 1974, 206 pp.

49. Yampol'skiy, S. M., Khilyuk, F. M., Lisichkin, V. A., "Problemy Nauchno-Tekhnicheskogo Prognozirovaniya" [Problems in Scientific and Technical Forecasting], Moscow, Ekonomika, 1969, 143 pp.

50. Yanch, E., "Prognozirovaniye Nauchno-Tekhnicheskogo Progressa" [Forecasting Scientific and Technical Progress], translated from the English, Progress, 1974, 586 pp.

10,272
CSO: 8144/1366

- END -